

Inhomogeneous random graphs, isolated vertices, and Poisson approximation

Mathew D. Penrose¹

University of Bath

July 28, 2015

Abstract

Consider a graph on randomly scattered points in an arbitrary space, with two points x, y connected with probability $\phi(x, y)$. Suppose the number of points is large but the mean number of isolated points is $O(1)$. We give general criteria for the latter to be approximately Poisson distributed. More generally, we consider the number of vertices of fixed degree, the number of components of fixed order, and the number of edges. We use a general result on Poisson approximation by Stein's method for a set of points selected from a Poisson point process. This method also gives a good Poisson approximation for Poisson U-statistics.

1 Introduction

In the *inhomogeneous random graph* (IRG), each vertex has one of several possible types or *states*, where the space of possible states may be infinite. Given the states of the vertices, each possible edge is present with a probability that depends on the states of the two endpoints, independently of the other edges. This provides a very flexible class of random graph models; for example, the state of a vertex could represent spatial location, or it could represent the time at which a vertex is born, for a randomly evolving graph. This class of models was introduced in [20], and has been studied in depth, for example in [4] and [2]. Much of the literature is concerned

¹ Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom: m.d.penrose@bath.ac.uk

Key words and phrases: Inhomogeneous random graph, random connection model, random geometric graph, Poisson approximation, Stein's method, U-statistic.

with the birth of the giant component, but full connectivity has also been studied in [6].

For our purposes, the IRG is defined as follows. Let $(\mathbb{X}, \mathcal{F}, \mu)$ be a probability space (the state space). For $s > 0$ let \mathcal{P}_s be the (random) set of points of a Poisson point process on \mathbb{X} with mean measure $s\mu$. Also, for $n \in \mathbb{N}$, let \mathcal{X}_n be the binomial point process consisting of n independent random elements of \mathbb{X} which common distribution μ .

Suppose $\phi : \mathbb{X} \times \mathbb{X} \rightarrow [0, 1]$ is a measurable symmetric function; we call such a function a *connection function*. Given finite $\mathcal{V} \subset \mathbb{X}$ (possibly with multiplicity), let $G(\mathcal{V}, \phi)$ be the random graph with vertex set \mathcal{V} , with each pair $\{x, y\}$ of points of \mathcal{V} connected by an edge with probability $\phi(x, y)$, independently of all other pairs. We are particularly interested in $G(\mathcal{P}_s, \phi)$ and $G(\mathcal{X}_n, \phi)$; we define these graphs more formally in Section 4.

In the special case where \mathbb{X} is a region of Euclidean space and $\phi(x, y)$ is determined by the displacement $x - y$ (typically via $\|x - y\|$, where $\|\cdot\|$ denotes the Euclidean norm), the IRG is also known as the *soft random geometric graph* [17] or *random connection model* (RCM) [15]. If, in the Euclidean setting, we have $\phi(x, y) = \mathbf{1}_{\{\|x-y\| \leq r\}}$ the IRG is known as the *random geometric graph* (RGG) [16] or *Gilbert graph*. These models are important in applications to wireless communications; see for example [7, 8, 10, 14, 21].

For any graph G and any $j \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, let $D_j(G)$ denote the number of vertices in G of degree j ; also set $D_{\leq j}(G) := \sum_{i=0}^j D_i(G)$. In particular, $D_0(G)$ is the number of isolated vertices. Of interest is the question of whether $D_{\leq k}(G(\mathcal{P}_s, \phi))$ and $D_{\leq k}(G(\mathcal{X}_{[s]}, \phi))$ are approximately Poisson distributed for s large, with k fixed, e.g. $k = 0$. One reason for interest is that $N_0(G) = 0$ (respectively $N_{\leq k}(G) = 0$) is clearly a necessary condition for G to be connected (resp. $(k+1)$ -connected), and for many choices of \mathbb{X} and ϕ_n this condition is asymptotically sufficient (in probability) when $G = G(\mathcal{X}_n, \phi_n)$ with n large; see [19, 17, 18, 11, 9]. In such cases, $\mathbb{P}[D_{\leq k}(G(\mathcal{X}_n, \phi_n)) = 0]$ is a good approximation for the probability that $G(\mathcal{X}_n, \phi_n)$ is $(k+1)$ -connected, so if we can estimate the former via Poisson approximation, we may have a useful estimate for the probability of $(k+1)$ -connectivity.

The references just mentioned show that in many cases where s is large and ϕ is chosen so that $\mathbb{E} D_j(G(\mathcal{P}_s, \phi)) = O(1)$, the distributions of $D_j(G(\mathcal{P}_s, \phi))$ and of $D_j(G(\mathcal{X}_{[s]}, \phi))$ are approximately Poisson. In this paper we give a general criterion for this to be the case, without making any geometrical or topological assumptions on the space \mathbb{X} (Theorem 2.1). We give a similar result for the number of components of order k with k fixed (Theorem 2.2), and for the number of edges (Theorem 2.3).

We prove Theorems 2.1, 2.2 and 2.3 in Sections 5, 6 and 7 respectively. The proofs use Theorem 3.1, a general result on Poisson approximation for functionals of Poisson processes, which is of independent interest. This theorem also gives us further results

on the Poisson approximation for the number of edges under different assumptions from those of Theorem 2.3. See Theorem 7.1 and the subsequent discussion. It also provides an alternative way to derive (and slightly improve) a result of [5] on Poisson approximation of Poisson U -statistics; see Section 8.

2 Results on inhomogeneous random graphs

Let Φ be the class of all measurable symmetric functions from $\mathbb{X} \times \mathbb{X}$ to $[0, 1]$, and for $\varepsilon > 0$ let Φ_ε be the class of all $\phi \in \Phi$ that satisfy

$$\inf_{x \in \mathbb{X}} \left(\int_{\mathbb{X}} \phi(x, y) \mu(dy) \right) \geq \varepsilon \sup_{x \in \mathbb{X}} \left(\int_{\mathbb{X}} \phi(x, y) \mu(dy) \right).$$

If $\phi \in \Phi_\varepsilon$ we say the connection function ϕ is ε -homogeneous. Note that 1-homogeneity, according to our definition, is the same as homogeneity as defined in [6].

Several of our results require ε -homogeneity. There are many interesting classes of connection function which satisfy this condition. For example, suppose \mathbb{X} is a bounded convex region in Euclidean space \mathbb{R}^d , and μ has a density (with respect to Lebesgue measure on that region) that is bounded away from 0 and infinity. Then there exists $\varepsilon > 0$ such that all connection functions of the form $\phi(x, y) = \psi(|x - y|)$ with ψ nonincreasing lie in Φ_ε . In particular, we do not require any exponential decay condition on ψ , such as that imposed in [17].

For $\phi \in \Phi$, set $\bar{\phi} = \sup_{x, y} \phi(x, y)$. Suppose we have connection functions ϕ_s defined for each $s > 0$ with $\mathbb{E} D_{\leq j}(G(\mathcal{P}_s, \phi_s))$ bounded as $s \rightarrow \infty$, for some fixed $j \in \mathbb{N}_0$. See (5.1) below for a formula for $\mathbb{E} D_j(G(\mathcal{P}_s, \phi_s))$. Our first result shows that $D_j(G(\mathcal{P}_s, \phi_s))$ is approximately Poisson for large s , under the condition that $\phi_s \in \Phi_\varepsilon$ for all s , along with an annoying extra condition that $\bar{\phi}_s$ is bounded away from 1.

For $\alpha \in (0, \infty)$, let Z_α denote a random variable having the Poisson distribution with parameter α . For any graph G we define $D_{\leq -1}(G) := 0$.

Theorem 2.1. *Let $j \in \mathbb{N}_0$. Suppose that for some $\varepsilon > 0$ and all $s \in (0, \infty)$ we have $\phi_s \in \Phi_\varepsilon$ and $\bar{\phi}_s \leq 1 - \varepsilon$. As $s \rightarrow \infty$, if $\mathbb{E} D_{\leq j}(G(\mathcal{P}_s, \phi_s)) \rightarrow \alpha \in (0, \infty)$, then*

$$D_j(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathcal{D}} Z_\alpha; \quad D_j(G(\mathcal{X}_{[s]}, \phi_{[s]})) \xrightarrow{\mathcal{D}} Z_\alpha, \quad (2.1)$$

and also

$$D_{\leq j-1}(G(\mathcal{P}_s, \phi_s)) \xrightarrow{P} 0; \quad D_{\leq j-1}(G(\mathcal{X}_{[s]}, \phi_{[s]})) \xrightarrow{P} 0. \quad (2.2)$$

It is interesting to compare the conclusion of this result with the example on page 55 of [16]. In that case, for a certain sequence of RGGs one can arrange for the number of vertices of degree 2 to be asymptotically *compound* Poisson, whereas here it is asymptotically Poisson.

For $k \in \mathbb{N}$, and for any graph G , we refer to the components of G of order k (i.e., with k vertices) as the k -components of G . Let $N_k(G)$ denote the number of k -components in G . In particular, $N_1(G) = D_0(G)$.

Suppose we have connection functions ϕ_s defined for each $s > 0$ with $\mathbb{E} N_k(G(\mathcal{P}_s, \phi_s))$ bounded as $s \rightarrow \infty$, for some fixed $k \in \mathbb{N}$. Our next result shows that $N_k(G(\mathcal{P}_s, \phi_s))$ and $N_k(G(\mathcal{X}_{[s]}, \phi_{[s]}))$ are approximately Poisson under the condition that $\phi_s \in \Phi_\varepsilon$ and $\bar{\phi}_s \leq 1 - \varepsilon$ for all s and some fixed $\varepsilon > 0$; even without this condition, the Poisson approximation holds provided that $\bar{\phi}_s = o(1/\log s)$ (in the case of $G(\mathcal{P}_s, \phi_s)$), or provided that $\bar{\phi}_{[s]} = o(s^{-1/2})$ (in the case of $G(\mathcal{X}_{[s]}, \phi_{[s]})$).

We also give a result on asymptotic normality, when $\mathbb{E} N_k(G(\mathcal{P}_s, \phi_s))$ grows slowly as $s \rightarrow \infty$. Let \mathcal{N} denote a random variable having the standard normal distribution in \mathbb{R} , i.e. with probability density function $(2\pi)^{-1} \exp(-x^2/2)$, $x \in \mathbb{R}$.

Theorem 2.2. *Suppose we are given $k \in \mathbb{N}, \alpha \in (0, \infty)$ and $\phi_s \in \Phi$ for $s > 0$.*

(a) Suppose $\mathbb{E} N_k(G(\mathcal{P}_s, \phi_s)) \rightarrow \alpha$ as $s \rightarrow \infty$. Then

(i) if $\bar{\phi}_s = o(1/\log s)$, then $N_k(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathcal{D}} Z_\alpha$ as $s \rightarrow \infty$;

(ii) if $\bar{\phi}_s = o(s^{-1/2})$ as $s \rightarrow \infty$, then $N_k(G(\mathcal{X}_{[s]}, \phi_{[s]})) \xrightarrow{\mathcal{D}} Z_\alpha$ as $s \rightarrow \infty$;

(iii) If there exists $\varepsilon > 0$ such that $\phi_s \in \Phi_\varepsilon$ and $\bar{\phi}_s \leq 1 - \varepsilon$ for all s , then as $s \rightarrow \infty$ we have

$$N_k(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathcal{D}} Z_\alpha; \quad N_k(G(\mathcal{X}_{[s]}, \phi_{[s]})) \xrightarrow{\mathcal{D}} Z_\alpha. \quad (2.3)$$

(b) Suppose instead as $s \rightarrow \infty$ that $\mathbb{E} N_k(G(\mathcal{P}_s, \phi_s)) \rightarrow \infty$, but $\mathbb{E} N_k(G(\mathcal{P}_s, \phi_s)) = o(s)$, and also that there exists $\varepsilon > 0$ such that $\phi_s \in \Phi_\varepsilon$ and $\bar{\phi}_s \leq 1 - \varepsilon$ for all s . Then setting $\alpha_s = \mathbb{E} N_k(G(\mathcal{P}_s, \phi_s))$, we have that $(N_k(G(\mathcal{P}_s, \phi_s)) - \alpha_s)/\sqrt{\alpha_s} \xrightarrow{\mathcal{D}} \mathcal{N}$.

Among other things, the case $k = 1$ of Theorem 2.2 (a) (i) generalizes Lemma 3.2 of [17] to a more general class of (\mathbb{X}, ϕ_s) than is considered in [17].

Our next result is concerned with Poisson or normal approximation for (a generalization of) the number of edges of $G(\mathcal{P}_s, \phi_s)$. This is of interest in itself, and will also be of use in the proof of Theorem 2.1. Given $k \in \mathbb{N}$ with $k \geq 2$, for any graph G let $H_k(G)$ denote the number of connected induced subgraphs of G of order k (so in particular, $H_2(G)$ is the number of edges.)

Theorem 2.3. *Let $\varepsilon > 0$ and $k \in \mathbb{N}$ with $k \geq 2$. Suppose $\phi_s \in \Phi_\varepsilon$ for all $s > 0$. Set $G_s = G(\mathcal{P}_s, \phi_s)$. Then:*

(i) As $s \rightarrow \infty$, if $\mathbb{E} H_k(G_s) \rightarrow \alpha \in (0, \infty)$, then $H_k(G_s) \xrightarrow{\mathcal{D}} Z_\alpha$.

(ii) As $s \rightarrow \infty$, if $\mathbb{E} H_k(G_s) \rightarrow \infty$ but $\mathbb{E} H_k(G_s) = o(s)$, then setting $\alpha_s := \mathbb{E} H_k(G_s)$, we have $(H_k(G_s) - \alpha_s)/\sqrt{\alpha_s} \xrightarrow{\mathcal{D}} \mathcal{N}$.

The proof of Theorems 2.1, 2.2 (iii) and 2.3 provides information about the rates of convergence in these results.

We now discuss some of the literature related to these results. Much of this concerns the RGG. Suppose $(\mathbb{X}, \mathcal{F}, \mu)$ is the d -dimensional Euclidean space with the Borel σ -algebra and with μ having a density f with respect to Lebesgue measure, while the connection function is given by $\phi_s(x, y) = \mathbf{1}_{\{|x-y| \leq r_s\}}$, with r_s chosen in such a way that $\mathbb{E} D_0(G(\mathcal{P}_s, \phi_s)) \rightarrow \alpha \in (0, \infty)$, i.e. (by (5.1) below)

$$s \int_{\mathbb{R}^d} \exp \left(-s \int_{B_{r_s}(x)} f(y) dy \right) f(x) dx \rightarrow \alpha \quad (2.4)$$

where $B_r(x) := \{y \in \mathbb{R}^d : |y - x| \leq r\}$.

Gupta and Iyer [9] consider the RGG over densities of the form $f(x) = c_1 \exp(-c_2|x|^{c_3})$, for r_s chosen to satisfy (2.4), and show that $D_0 \xrightarrow{\mathcal{D}} Z_\alpha$ in this case. This generalises earlier work in [18] on the case where $c_3 = 2$ (i.e., the standard normal distribution).

Hsing and Rootzén [11] consider the RGG in $d = 2$ for a general class of densities with unbounded support. In their Theorem 1, they give a general set of conditions, including (2.4), under which $D_0(G(\mathcal{P}_s, r_s)) \xrightarrow{\mathcal{D}} Z_\alpha$. They show that these conditions apply, for example when the density f is bivariate normal with standardized marginals, or has independent Weibull distributed coordinates.

Suppose $(\mathbb{X}, \mathcal{F}, \mu)$ is the d -dimensional unit cube equipped with Lebesgue measure. Suppose ϕ_s is of the form $\phi_s(x, y) = p_s \mathbf{1}_{\{|x-y| \leq r_s\}}$, with p_s and r_s chosen so that $\mathbb{E} D_0(G(\mathcal{P}_s, \phi_s)) \rightarrow \alpha$, then $D_0 \xrightarrow{\mathcal{D}} Z_\alpha$. More generally, if ϕ_s is of the form $\phi_s(x, y) = \psi_s(|x - y|)$ with ψ_s a decreasing function satisfying an exponential decay condition, then the Poisson convergence of D_0 is known to hold; see [17]. Our results enable us to relax the exponential decay condition and allow for other distributions, for example with density bounded away from zero and infinity on a convex compact region in \mathbb{R}^d . In the special case where $\phi_s(x, y) = \mathbf{1}_{\{|x-y| \leq r_s\}}$, several results similar to those in our presented here are given in [16]; see Theorems 3.4 and 8.1 there.

It is tempting to think that for α fixed and s large, for *any* choice of connection function ϕ satisfying $\mathbb{E} D_0(G(\mathcal{P}_s, \phi)) = \alpha$, the distribution of $D_0(G(\mathcal{P}_s, \phi))$ would be approximately Z_α . For example, if $\phi(x, y) \in \{0, 1\}$ for all x, y , then intuition might suggest that for large s , if ϕ is tuned so that $\mathbb{E} D_0(G(\mathcal{P}_s, \phi)) = \alpha$, then each isolated vertex is a ‘rare event’, and the information that a particular vertex (at x , say) is isolated affects only its neighbourhood (i.e. the region $\{y \in \mathbb{X} : \phi(x, y) = 1\}$).

However, this is not always the case. See the last paragraph of [11, Section 2]. Suppose for example that $(\mathbb{X}, \mathcal{F}, \mu)$ is the unit interval equipped with Lebesgue

measure, $\phi_s(x, y)$ is equal to 1 if $\max(x, y) \leq s^{-1}$ or $\min(x, y) > s^{-1}$, and otherwise $\phi_s(x, y) = 0$. Then in the large- s limit the random variable $D_0(G(\mathcal{P}_s, \phi_s))$ is Bernoulli distributed with parameter e^{-1} , not Poisson distributed. In this case the event that there is an isolated vertex in $[0, 1/s]$ is not rare. The condition of ε -homogeneity which appears in many of our results, rules out this sort of example.

Devroye and Fraiman [6] consider $D_0(G(\mathcal{X}_n, \phi_n))$ on a general space \mathbb{X} in the case where $\phi_n = (a(n)\kappa) \wedge 1$ for a fixed function $\kappa : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}_+$ (such a function κ is called a *kernel*), and some sequence $a(n)$ (where \wedge denotes minimum). This is a common assumption for the IRG; see for example [4]. Their results may be interpreted as saying that (under certain conditions) the threshold value of a above which $G(\mathcal{X}_n, a\kappa \wedge 1)$ is free of isolated vertices (which may be viewed as a random variable) satisfies a weak law of large numbers; it is asymptotic to a constant times $(\log n)/n$, in probability. They also derive a similar law of large numbers for the threshold for the graph to be connected.

When κ is bounded, our Theorem 2.2 (part (a) (ii)) gives a possibility of deriving convergence in distribution for this threshold (suitably transformed). For example, suppose $\mathbb{X} = [0, 1)$ and μ has a density f with respect to Lebesgue measure, with f bounded away from zero. Suppose also that $\kappa(x, y)$ is determined via the absolute value of $x - y \pmod{1}$ (so we are in the one-dimensional torus). If $f \equiv 1$, or if f is smooth with $g(x) := \int \kappa(x, y)f(y)dy$ having a unique minimum, then it should be possible to derive a distributional limit law for this threshold, since in these cases it should be possible, for any $\alpha \in (0, \infty)$, to determine a sequence a_s such that $\mathbb{E} N_0(G(\mathcal{P}_s, a_s\kappa)) \rightarrow \alpha$, and then apply Theorem 2.2 (a) (ii).

If κ is unbounded but shift-invariant on the torus, for example if $\kappa(x, y) = ((x - y) \bmod 1)^{-\gamma}$ for some fixed positive γ and $f \equiv 1$, then for any $\varepsilon \in (0, 1)$, it may be possible to find a limiting distribution for a suitable transformation of the threshold value of a above which $G(\mathcal{X}_n, a\kappa \wedge (1 - \varepsilon))$ is free of isolated vertices, now using Theorem 2.2 (a) (iii). Indeed, in this case the connection functions are all 1-homogeneous, and by using connection functions $\phi_n = a_n\kappa \wedge (1 - \varepsilon)$ rather than the more standard $\phi_n = a_n\kappa \wedge 1$, we ensure that the extra condition $\overline{\phi}_n \leq 1 - \varepsilon$ is also satisfied.

It would be interesting to fully work out and extend these examples; to get a similar results for the connectivity threshold; and to improve the weak law of [6] to a strong law.

3 A general result on Poisson approximation

Let $(\mathbb{M}, \mathcal{M}, \mathbf{m})$ be a probability space (known as a *mark space*). Assume that the probability measure \mathbf{m} on \mathbb{M} is *diffuse*, by which we mean that there is a product

measurable set $A \subset \mathbb{M} \times \mathbb{M}$ with $(\mathbf{m} \otimes \mathbf{m})(A) = 0$, such that the diagonal $\{(t, t) : t \in \mathbb{M}\}$ is contained in A . For example, if $\{t\} \in \mathcal{M}$ and $\mathbf{m}(\{t\}) = 0$ for all $t \in \mathbb{M}$, then \mathbf{m} is diffuse.

Suppose on a suitable probability space that we have a sequence $((X_i, T_i), i = 1, 2, 3, \dots)$ of independent identically distributed random elements of $\mathbb{X} \times \mathbb{M}$ with common distribution $\mu \otimes \mathbf{m}$, and an independent unit rate Poisson counting process $(Z_s, s > 0)$, so that the random variable Z_s has $\text{Poisson}(s)$ distribution for each s , and also a further independent sequence $(\tau, \tau_1, \tau_2, \tau_3, \dots)$ of independent random elements of \mathbb{M} with common distribution \mathbf{m} . By our assumption that \mathbf{m} is diffuse, the values of $\tau, T_1, \tau_1, T_2, \tau_2, T_3, \tau_3, \dots$ are almost surely distinct.

A *finite point process* in \mathbb{X} is defined as a random element of the space $\mathbf{S}(\mathbb{X})$ of all finite subsets of \mathbb{X} equipped with the smallest σ -algebra $\mathcal{S}(\mathbb{X})$ containing the sets $\{\xi \in \mathbf{S}(\mathbb{X}) : \xi(B) = k\}$ for all $B \in \mathcal{F}$ and all $k \in \mathbb{N}_0$, where $\xi(B) := |\xi \cap B|$ and $|\xi|$ denotes the number of elements of ξ .

A finite point process in $\mathbb{X} \times \mathbb{M}$ is defined similarly as a random element of the space $\mathbf{S} := \mathbf{S}(\mathbb{X} \times \mathbb{M})$, where $\mathbb{X} \times \mathbb{M}$ is equipped with the product σ -algebra $\mathcal{F} \otimes \mathcal{M}$. For $k \in \mathbb{N}$, let $\mathbf{S}_k := \{\xi \in \mathbf{S} : |\xi| = k\}$.

Let $s > 0, n \in \mathbb{N}$. Define the following point processes in $\mathbb{X} \times \mathbb{M}$:

$$\eta := \cup_{i=1}^{Z_s} \{(X_i, T_i)\}; \quad \xi_n := \cup_{i=1}^n \{(X_i, T_i)\}. \quad (3.1)$$

Then η is a Poisson point process in $\mathbb{X} \times \mathbb{M}$ with mean measure $\lambda \times \mathbf{m}$, where we set $\lambda = s\mu$. Similarly, ξ_n is a binomial point process in $\mathbb{X} \times \mathbb{M}$.

Let d_{TV} and d_W denote total variation distance and Wasserstein distance, respectively, between probability measures on the nonnegative integers. That is, for \mathbb{N}_0 -valued random variables X, Y with distribution $\mathcal{L}(X), \mathcal{L}(Y)$ respectively, we set

$$d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{A \subset \mathbb{N}_0} (\mathbb{P}[X \in A] - \mathbb{P}[Y \in A]);$$

$$d_W(\mathcal{L}(X), \mathcal{L}(Y)) = \sup\{\|\mathbb{E} h(X) - \mathbb{E} h(Y)\| : \|\Delta h\|_\infty \leq 1\},$$

where for $h : \mathbb{N}_0 \rightarrow \mathbb{R}$ we set $\Delta h(i) := h(i+1) - h(i)$ for $i \in \mathbb{N}_0$, and $\|h\|_\infty := \sup_{i \in \mathbb{N}_0} |h(i)|$.

The following theorem is related to a well-known result on the Poisson approximation of a sum of Bernoulli random variables by Stein's method via coupling (Theorem II.24.3 of [13], or Theorem 1.B of [1]). Here the terms in the sum are themselves indexed by k -subsets of the set of points of a (marked) Poisson point process.

Let $k \in \mathbb{N}$ and let $f : \mathbf{S}_k \times \mathbf{S} \rightarrow \{0, 1\}$ be a measurable function. For $\xi \in \mathbf{S}$, set

$$F(\xi) := \sum_{\{\psi \subset \xi : |\psi| = k\}} f(\psi, \xi \setminus \psi) \quad (3.2)$$

We can think of f as a mechanism for selecting some of the k -subsets of ξ , and $F(\xi)$ as the total number of k -subsets selected.

For $x_1, \dots, x_k \in \mathbb{X}$ set $p(x_1, \dots, x_k) := \mathbb{E} f(\{(x_1, \tau_1), \dots, (x_k, \tau_k)\}, \eta)$.

Theorem 3.1. *Let $W := F(\eta)$ with η and F as described above. Suppose that $w : \mathbb{X}^k \rightarrow [0, \infty)$ is a measurable function, and that for λ^k -almost every $\mathbf{x} = (x_1, \dots, x_k) \in \mathbb{X}^k$ with $p(x_1, \dots, x_k) > 0$ we can find coupled random variables $U_{\mathbf{x}}, V_{\mathbf{x}}$ such that*

- $\mathcal{L}(U_{\mathbf{x}}) = \mathcal{L}(W)$;
- $\mathcal{L}(1 + V_{\mathbf{x}}) = \mathcal{L}(F(\cup_{i=1}^k \{(x_i, \tau_i)\} \cup \eta) | f(\{(x_1, \tau_1), \dots, (x_k, \tau_k)\}, \eta) = 1)$.
- $\mathbb{E} [|U_{\mathbf{x}} - V_{\mathbf{x}}|] \leq w(\mathbf{x})$.

Set $\alpha = \mathbb{E} W = (1/k!) \int p(\mathbf{x}) \lambda^k(d\mathbf{x})$. Then

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(Z_{\alpha})) \leq \frac{(1 \wedge \alpha^{-1})}{k!} \int w(\mathbf{x}) p(\mathbf{x}) \lambda^k(d\mathbf{x}), \quad (3.3)$$

and

$$d_W(\mathcal{L}(W), \mathcal{L}(Z_{\alpha})) \leq \frac{3(1 \wedge \alpha^{-1/2})}{k!} \int w(\mathbf{x}) p(\mathbf{x}) \lambda^k(d\mathbf{x}). \quad (3.4)$$

The proof of this uses the following fact. Suppose $g : \mathbf{S}_k \times \mathbf{S} \rightarrow \mathbb{R}$ is a bounded measurable function. Then

$$\mathbb{E} \sum_{\{\psi \subset \eta; |\psi|=k\}} g(\psi, \eta \setminus \psi) = \frac{1}{k!} \int \mathbb{E} [g(\{(x_1, \tau_1), \dots, (x_k, \tau_k)\}, \eta)] \lambda^k(d\mathbf{x}). \quad (3.5)$$

This is known as the (multivariate) *Mecke formula*. See e.g. [16, Theorem 1.6]. This fact gives us the assertion in the statement of the theorem that $\mathbb{E} W = (1/k!) \int p(x) \lambda^k(dx)$.

Theorem 3.1 still holds in the case where the measure λ is σ -finite but infinite, so that the Poisson point process η with mean measure $\lambda \otimes \mathbf{m}$ is almost surely infinite. The proof is essentially unchanged.

Proof of Theorem 3.1. Let $h : \mathbb{N}_0 \rightarrow \mathbb{R}$ be bounded. Then

$$\begin{aligned} \mathbb{E} [Wh(W)] &= \mathbb{E} \sum_{\{\psi \subset \eta; |\psi|=k\}} f(\psi, \eta \setminus \psi) h(F(\eta)) \\ &= \frac{1}{k!} \int \mathbb{E} [f(\{(x_1, \tau_1), \dots, (x_k, \tau_k)\}, \eta) h(F(\cup_{i=1}^k \{(x_i, \tau_i)\} \cup \eta))] \lambda^k(d\mathbf{x}) \\ &= \frac{1}{k!} \int \mathbb{E} [h(F(\cup_{i=1}^k \{(x_i, \tau_i)\} \cup \eta)) | f(\cup_{i=1}^k \{(x_i, \tau_i)\}, \eta) = 1] p(\mathbf{x}) \lambda(d\mathbf{x}). \end{aligned}$$

Also, $\mathbb{E}[\alpha h(W + 1)] = (1/k!) \int \mathbb{E} h(W + 1) p(\mathbf{x}) \lambda^k(d\mathbf{x})$, and therefore

$$\begin{aligned} |\mathbb{E}[\alpha h(W + 1) - Wh(W)]| &\leq \frac{1}{k!} \int p(\mathbf{x}) \lambda(d\mathbf{x}) \\ &\times \left| \mathbb{E} h(W + 1) - \mathbb{E} \left[h \left(F \left(\bigcup_{i=1}^k \{(x_i, \tau_i)\} \cup \eta \right) \right) \middle| f \left(\bigcup_{i=1}^k \{(x_i, \tau_i)\}, \eta \right) = 1 \right] \right|. \end{aligned}$$

Since $|h(i) - h(j)| \leq \|\Delta h\|_\infty \cdot |i - j|$ for $i, j \in \mathbb{N}_0$, we obtain for each $\mathbf{x} = (x_1, \dots, x_k) \in \mathbb{X}^k$ that

$$\begin{aligned} |\mathbb{E} h(W + 1) - \mathbb{E} \left[h \left(F \left(\bigcup_{i=1}^k \{(x_i, \tau_i)\} \cup \eta \right) \right) \middle| f \left(\bigcup_{i=1}^k \{(x_i, \tau_i)\}, \eta \right) = 1 \right]| \\ \leq |\mathbb{E} h(U_{\mathbf{x}} + 1) - \mathbb{E} h(V_{\mathbf{x}} + 1)| \\ \leq \|\Delta h\|_\infty \mathbb{E} |U_{\mathbf{x}} - V_{\mathbf{x}}| \leq \|\Delta h\|_\infty w(\mathbf{x}), \end{aligned}$$

and therefore

$$|\mathbb{E}[\alpha h(W + 1) - Wh(W)]| \leq \frac{\|\Delta h\|_\infty}{k!} \int w(\mathbf{x}) p(\mathbf{x}) \lambda^k(d\mathbf{x}).$$

Given $A \subset \mathbb{N}_0$, set $g = \mathbf{1}_A$ and choose $h : \mathbb{N}_0 \rightarrow \mathbb{R}$ so that $h(0) = 0$ and

$$\alpha h(i + 1) - ih(i) = g(i) - Z_\alpha(g), i \in \mathbb{N}_0. \quad (3.6)$$

Then (see Lemma 1.1.1 of [1]) h is bounded and $\|\Delta h\|_\infty \leq 1 \wedge \alpha^{-1}$, and hence

$$|\mathbb{P}[W \in A] - Z_\alpha(A)| \leq \frac{(1 \wedge \alpha^{-1})}{k!} \int w(\mathbf{x}) p(\mathbf{x}) \lambda^k(d\mathbf{x}).$$

The result (3.3) follows.

One obtains (3.4) similarly by choosing, for any given g with $|g(i) - g(i - 1)| \leq 1$ for $i \in \mathbb{N}$, a solution h to (3.6) with $h(0) = 0$, and using Lemma 1.1.5 of [1]. \square

4 Formal constructions of the IRG

Let $s \in (0, \infty)$ and $n \in \mathbb{N}$. Let $\phi \in \Phi$. We now give a more formal definition of the graphs $G(\mathcal{P}_s, \phi)$ and $G(\mathcal{X}_n, \phi)$,

We make the following particular choice of mark space $(\mathbb{M}^*, \mathcal{M}^*, \mathbf{m}^*)$. Let Leb denote Lebesgue measure on $[0, 1)$. Let $\mathbb{M}^* = [0, 1)^{\mathbb{N}_0}$ with \mathcal{M}^* the product Borel σ -algebra, and with $\mathbf{m}^* := \otimes_{n=0}^\infty \text{Leb}$, so that a random element of \mathbb{M}^* with distribution \mathbf{m}^* is a sequence of independent $\text{unif}(0, 1)$ random variables indexed by \mathbb{N}_0 .

Now (and for the rest of this paper) taking $(\mathbb{M}, \mathcal{M}, \mathbf{m}) = (\mathbb{M}^*, \mathcal{M}^*, \mathbf{m}^*)$, let $(X_i, T_i)_{i \geq 1}$, $(Z_s)_{s > 0}$ and $(\tau, \tau_1, \tau_2, \dots)$ be as in the preceding section. Given $n \in \mathbb{N}$ and $s > 0$, let the point processes η and ξ_n be as given by (3.1). We now write η_s for η

to emphasise the dependence on s . Thus η_s is a Poisson point process in \mathbb{X} with mean measure $s\mu$ and with each point marked with a sequence of independent uniform $[0, 1]$ variables indexed by the nonnegative integers. Similarly, ξ_n is a binomial point process in $\mathbb{X} \times \mathbb{M}^*$. We write \mathbf{S}^* for $\mathbf{S}(\mathbb{X} \times \mathbb{M}^*)$.

To ease notation, we shall also assume from now on that the probability measure μ on \mathbb{X} is *diffuse*. This ensures that the values of X_1, X_2, \dots are almost surely distinct. However, this assumption is for notational convenience only; even without it, one can make sense of our results either by allowing the set $\{X_1, \dots, X_n\}$ to have multiplicities, or by using the attached marks T_i (which are almost surely distinct) to distinguish between different points X_i .

Set $\mathcal{X}_n := \{X_1, \dots, X_n\}$ and $\mathcal{P}_s := \{X_1, \dots, X_{Z_s}\}$, the canonical projections of ξ_n and η_s respectively onto \mathbb{X} . Define the graph $G(\mathcal{X}_n, \phi)$ to have vertex set \mathcal{X}_n and to have an edge between vertices X_i and X_j , for $i, j \in [n] := \{1, 2, \dots, n\}$ with $i < j$, if and only if $T_{i,j} \leq \phi(X_i, X_j)$, where $T_i = (T_{i,0}, T_{i,1}, T_{i,2}, \dots)$. Let $G(\mathcal{P}_s, \phi)$ be the graph $G(\mathcal{X}_{Z_s}, \phi)$.

This is one way to formally define the random graphs with the properties described more informally in the Introduction. It has the advantage that $G(\mathcal{X}_n, \phi)$ is a subgraph of $G(\mathcal{P}_s, \phi)$ whenever $n \leq Z_s$ and $G(\mathcal{P}_s, \phi)$ is a subgraph of $G(\mathcal{X}_n, \phi)$ whenever $Z_s \leq n$, which is useful for coupling arguments. However, it has the disadvantage that the edge-set of the graph $G(\mathcal{P}_s, \phi)$ is not invariant under permutation of the order in which the marked points $(X_1, T_1), \dots, (X_{Z_s}, T_{Z_s})$ are listed. Therefore we define a further graph which has the same distribution but also satisfies this permutation-invariance. This will be useful in applying Theorem 3.1 in the proof of Theorems 2.1, 2.2 and 2.3.

Given $\xi \in \mathbf{S}^*$ and $\phi \in \Phi$, define the graph $G_\phi(\xi)$ as follows. If there exist distinct $(x, t_0, t_1, \dots) \in \xi$ and $(y, u_0, u_1, \dots) \in \xi$ with $t_0 = u_0$, then set $G_\phi(\xi)$ to be the empty graph. Otherwise, we can write ξ uniquely as

$$\xi = \bigcup_{i=1}^{|\xi|} \{(x_i, t_{i,0}, t_{i,1}, t_{i,2}, \dots)\}$$

with $t_{1,0} < t_{2,0} < \dots < t_{|\xi|,0}$. Let $G_\phi(\xi)$ have vertex set $\{x_1, \dots, x_{N(\xi)}\}$, and for each $i < j \leq |\xi|$ let $G_\phi(\xi)$ have an edge connecting x_i to x_j if and only if $t_{i,j} \leq \phi(x_i, x_j)$. In other words, we use the first coordinate of the marks to determine the order in which we enumerate the points of ξ ; having done so, for $i < j$ we use the $(j+1)$ -st component of the mark attached to the i -th point to decide whether to connect it to the j -th point.

Let us say that two random graphs G and G' have the same distribution if any graph invariant of G has the same distribution as the same graph invariant evaluated on G' . The following is immediate from the independence of the components of the marks τ_i .

Proposition 4.1. *For any finite $\mathcal{X} = \{x_1, \dots, x_m\} \subset \mathbb{X}$ the distribution of the random graph $G_\phi(\{(x_1, \tau_1), \dots, (x_m, \tau_m)\})$ is the same as that of $G(\mathcal{X}, \phi)$.*

In particular, the the distribution of $G_\phi(\eta_s)$ is the same as that of $G(\mathcal{P}_s, \phi)$, although they are not the same graph because the set of edges is defined differently for the two graphs. Likewise $G_\phi(\xi_n)$ has the same distribution as $G(\mathcal{X}_n, \phi)$.

In the following lemmas we check measurability of functions which will feature in the proof of Theorems 2.1 and 2.2 (a) (iii) respectively. We use the following notation. For $\ell \in \mathbb{N}$ and $i \in [\ell]$, let $I_{\ell,i}$ denote the set of $(x, t_0, t_1, t_2, \dots) \in \mathbb{X} \times [0, 1]^{\mathbb{N}_0}$ such that $(i-1)/\ell \leq t_0 < i/\ell$.

Lemma 4.1. *Let $j \in \mathbb{N}_0$. Then the function $f : \mathbb{X} \times \mathbb{M}^* \times \mathbf{S}^* \rightarrow \{0, 1\}$ given by*

$$f(x, \mathbf{t}, \xi) := \mathbf{1}\{x \text{ has degree } j \text{ in } G_\phi(\xi \cup \{(x, \mathbf{t})\})\}, \quad (x, \mathbf{t}, \xi) \in \mathbb{X} \times \mathbb{M}^* \times \mathbf{S}^*$$

is measurable.

Proof. For $\ell, n, m, i_1, \dots, i_n \in \mathbb{N}$ with $m \leq n \leq \ell$, and $i_1 < i_2 < \dots < i_n \leq \ell$, let $A_{\ell,n,m,i_1,\dots,i_n}$ denote the set of $(x, \mathbf{t}, \xi) \in \mathbb{X} \times \mathbb{M}^* \times \mathbf{S}^*$ such that:

1. $(\xi \cup \{(x, \mathbf{t})\})(I_{\ell,i_k}) = 1$, for $1 \leq k \leq n$;
2. $(\xi \cup \{(x, \mathbf{t})\})(I_{\ell,i}) = 0$, for $i \in [\ell] \setminus \{i_1, \dots, i_n\}$;
3. $(x, \mathbf{t}) \in I_{\ell,i_m}$;
4. $(\sum_{k=1}^{m-1} \xi(\{(y, \mathbf{u}) \in I_{\ell,i_k} : u_m \leq \phi(x, y)\})) + \sum_{k=m+1}^n \xi(\{(y, \mathbf{u}) \in I_{\ell,i_k} : t_k \leq \phi(x, y)\}) = j$, where $\mathbf{u} = (u_0, u_1, \dots)$ and $\mathbf{t} = (t_0, t_1, \dots)$.

Then each $A_{\ell,n,m,i_1,\dots,i_n}$ is measurable in $\mathbb{X} \times \mathbb{M}^* \times \mathbf{S}^*$, and

$$f(x, \mathbf{t}, \xi) = \mathbf{1}\{(x, \mathbf{t}, \xi) \in \cup_{\ell,n,m,i_1,\dots,i_n \in \mathbb{N}: m \leq n \leq \ell, 1 \leq i_1 < i_2 < \dots < i_n \leq \ell} A_{\ell,n,m,i_1,\dots,i_n}\}$$

which is a measurable function. □

Lemma 4.2. *Let $k \in \mathbb{N}$. Suppose the function $\tilde{f} : (\mathbb{X} \times \mathbb{M}^*)^k \times \mathbf{S}^* \rightarrow \{0, 1\}$ is given, for $(x_1, \mathbf{t}_1, \dots, x_k, \mathbf{t}_k) \in (\mathbb{X} \times \mathbb{M}^*)^k$ and $\xi \in \mathbf{S}^*$, by*

$$\begin{aligned} & \tilde{f}(x_1, \mathbf{t}_1, \dots, x_k, \mathbf{t}_k, \xi) \\ & := \mathbf{1}\{\{x_1, \dots, x_k\} \text{ induces a component of } G_\phi(\cup_{i=1}^k \{(x_i, \mathbf{t}_i)\} \cup \xi)\}. \end{aligned}$$

Then \tilde{f} is measurable.

Proof. For $\ell, n, m_1, \dots, m_k, i_1, \dots, i_n \in \mathbb{N}$ with $\max_{1 \leq i \leq k} m_i \leq n \leq \ell$, and m_1, \dots, m_k distinct, and $1 \leq i_1 < i_2 < \dots < i_n \leq \ell$, let $A_{\ell, n, m_1, \dots, m_k, i_1, \dots, i_n}$ denote the set of $(x_1, \mathbf{t}_1, \dots, x_k, \mathbf{t}_k, \xi) \in (\mathbb{X} \times \mathbb{M}^*)^k \times \mathbf{S}^*$ such that for some connected graph Γ on vertex set $[k]$:

1. $(\cup_{h=1}^k \{(x_h, \mathbf{t}_h)\} \cup \xi) (I_{\ell, i_j}) = 1$, for $1 \leq j \leq n$;
2. $(\cup_{h=1}^k \{(x_h, \mathbf{t}_h)\} \cup \xi) (I_{\ell, i}) = 0$, for $i \in [\ell] \setminus \{i_1, \dots, i_n\}$;
3. $(x_h, \mathbf{t}_h) \in I_{\ell, i_{m_h}}$ for $1 \leq h \leq k$;
4. for all $h \in [k]$ and $1 \leq j < m_h$ with $j \notin \{m_1, \dots, m_k\}$ we have $\xi(\{(y, \mathbf{u}) \in I_{i_j} : u_{m_h} > \phi(x_h, y)\}) = 1$, where $\mathbf{u} = (u_0, u_1, \dots)$;
5. for all $h \in [k]$ and $m_h < j \leq n$ with $j \notin \{m_1, \dots, m_k\}$ we have $\xi(\{(y, \mathbf{u}) \in I_{i_j} : t_{h,j} > \phi(x_h, y)\}) = 1$, where $\mathbf{t}_h = (t_{h,0}, t_{h,1}, \dots)$, and
6. for all $h, h' \in [k]$ with $m_h < m_{h'}$ and $\{h, h'\}$ an edge of Γ , we have $t_{h, m_h} \leq \phi(x_h, x_{h'})$.

Then each $A_{\ell, n, k, m_1, \dots, m_k, i_1, \dots, i_n}$ is measurable in $(\mathbb{X} \times \mathbb{M})^k \times \mathbf{S}^*$, and setting $[n]_{\neq}^k$ to be the set of $(m_1, \dots, m_k) \in [n]^k$ such that m_1, \dots, m_k are distinct, we have that \tilde{f} is the indicator of the set

$$\cup_{\ell, n, i_1, \dots, i_n \in \mathbb{N}, (m_1, \dots, m_k) \in [n]_{\neq}^k : n \leq \ell, 1 \leq i_1 < i_2 < \dots < i_n \leq \ell} A_{\ell, n, m_1, \dots, m_k, i_1, \dots, i_n}$$

which is measurable. □

5 Proof of Theorem 2.1

For $s > 0$, let $\eta_s, Z_s, \tau, \tau_1, \tau_2, \dots$ be as in the preceding section. For $\phi \in \Phi$ and $i \in \mathbb{N}_0$, we have by the Mecke equation (3.5) that

$$\mathbb{E} D_i(G_\phi(\eta_s)) = s \int_{\mathbb{X}} \frac{(s \int \phi(x, y) \mu(dy))^i}{i!} \exp \left(-s \int \phi(x, y) \mu(dy) \right) \mu(dx). \quad (5.1)$$

In the sequel, other formulae for expectations of numbers of vertices, or k -tuples of vertices, having certain properties in terms of the graph $G_\phi(\eta_s)$ (or equivalently, the graph $G(\mathcal{P}_s, \phi)$), will also be justified by the Mecke formula. These arguments also justify (3.5) of [17].

Proof of Theorem 2.1. Let $i \in \mathbb{N}_0$. For $s > 0$, set $a_s = \sup_{x \in \mathbb{X}} \int \phi_s(x, y) \mu(dy)$. Since we assume $\phi_s \in \Phi_\varepsilon$ for all $s > 0$, we have for $s > 0$ and $i \in \mathbb{N}$ that

$$\frac{\varepsilon s a_s}{i} \leq \frac{\mathbb{E} D_i(G(\mathcal{P}_s, \phi_s))}{\mathbb{E} D_{i-1}(G(\mathcal{P}_s, \phi_s))} \leq \frac{s a_s}{i}. \quad (5.2)$$

Now fix $j \in \mathbb{N}_0$ and suppose ϕ_s is such that as $s \rightarrow \infty$ we have

$$\mathbb{E} D_{\leq j}(G(\mathcal{P}_s, \phi_s)) \rightarrow \alpha \in (0, \infty). \quad (5.3)$$

Suppose that $s a_s$ remains bounded as $s \rightarrow \infty$ along some subsequence. Then by (5.1), $\mathbb{E} D_0(G(\mathcal{P}_s, \phi_s)) \rightarrow \infty$ along that subsequence, which contradicts (5.3). Therefore $s a_s \rightarrow \infty$ as $s \rightarrow \infty$.

By (5.3), $\mathbb{E} D_j(G(\mathcal{P}_s, \phi_s))$ remains bounded, and since also $s a_s \rightarrow \infty$, if $j \geq 1$ then using (5.2) we have that $\mathbb{E} D_{j-1}(G(\mathcal{P}_s, \phi_s)) \rightarrow 0$, and repeating the argument we also have

$$\mathbb{E} D_i(G(\mathcal{P}_s, \phi_s)) \rightarrow 0, \quad i = 0, 1, \dots, j-1; \quad (5.4)$$

$$\mathbb{E} D_j(G(\mathcal{P}_s, \phi_s)) \rightarrow \alpha. \quad (5.5)$$

The first part of (2.2) follows from (5.4).

For any two $(0, \infty)$ -valued functions $u(s)$ and $v(s)$, we say $u(s) = \Theta(v(s))$ if $\liminf_{s \rightarrow \infty} u(s)/v(s) > 0$ and $\limsup_{s \rightarrow \infty} u(s)/v(s) < \infty$. We show next that

$$s a_s = \Theta(\log s). \quad (5.6)$$

We already know that $s a_s \rightarrow \infty$. Since for any $i \in \mathbb{N}$ the function $x^i e^{-x}$ is decreasing in x for large enough x , for large enough s we have that if $s a_s < (\log s)/2$ then by (5.1),

$$\mathbb{E} D_j(G(\mathcal{P}_s, \phi_s)) \geq (s/j!) e^{-\log s/2} ((\log s)/2)^j,$$

which tends to infinity, contradicting (5.3). On the other hand, since $\phi_s \in \Phi_\varepsilon$, if $s a_s > (2/\varepsilon) \log s$ then for large enough s , by (5.1) we have

$$\mathbb{E} D_j(G(\mathcal{P}_s, \phi_s)) \leq s e^{-\varepsilon(2/\varepsilon) \log s} ((2/\varepsilon) \log s)^j,$$

which tends to zero, contradicting (5.5). Thus we have (5.6).

Next we show that $D_j(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathcal{D}} Z_\alpha$. We aim to apply Theorem 3.1 to the case $\phi = \phi_s$ of the function f considered in Lemma 4.1. That is, for $(x, \mathbf{t}, \xi) \in \mathbb{X} \times \mathbb{M}^* \times \mathbf{S}^*$ we set $f_s(x, \mathbf{t}, \xi)$ to be the indicator of the statement that x has degree j in $G_{\phi_s}(\xi \cup \{(x, \mathbf{t})\})$. Then $D_j(G_{\phi_s}(\eta_s)) = F_s(\eta_s)$, where F_s is the function F obtained by using $f \equiv f_s$ in (3.2) (with $k = 1$).

Let $s > 0$ and $x \in \mathbb{X}$. If $j \geq 1$ suppose we also have an extra sequence (Y, Y_1, \dots, Y_j) of independent identically distributed random elements of \mathbb{X} with $\mathbb{P}[Y \in dy] = \phi_s(x, y)\mu(dy) / \int \phi_s(x, z)\mu(dz)$, independent of $((X_i, T_i))_{i \geq 1}$ and $(Z_s)_{s > 0}$ and $(\tau, \tau_1, \tau_2, \dots)$. Let \mathcal{G}_s be the graph $G_{\phi_s}(\eta_s \cup \{(x, \tau), (Y_1, \tau_1), \dots, (Y_k, \tau_k)\})$, with added edges from x to each of Y_1, \dots, Y_k (if not already included). Let $\mathcal{P}_{s,x}$ be the set of points of \mathcal{P}_s that are connected to x in this graph, and set $\mathcal{P}_s^x := \mathcal{P}_s \setminus \mathcal{P}_{s,x}$.

Let U_x denote the number of vertices of degree j in the subgraph of \mathcal{G}_s induced by vertex set \mathcal{P}_s . By Proposition 4.1, this graph has the distribution of $G(\mathcal{P}_s, \phi_s)$, so U_x has the distribution of $F_s(\eta_s)$.

Now consider the subgraph of \mathcal{G}_s induced by $\{x\} \cup \mathcal{P}_s$. This has the distribution of $G(\mathcal{P}_s \cup \{x\}, \phi_s)$, and $\mathcal{P}_{s,x}$ is the set of vertices in this graph lying adjacent to x . Conditioning on x having degree j amounts to conditioning on $|\mathcal{P}_{s,x}| = j$. We define a coupled point process \mathcal{P}^* (a subset of $\mathcal{P}_s \cup \{Y_1, \dots, Y_j\}$) with the distribution of \mathcal{P}_s conditioned on x having degree j , as follows.

If $|\mathcal{P}_{s,x}| > j$ then we select $|\mathcal{P}_{s,x}| - j$ elements of $\mathcal{P}_{s,x}$ uniformly at random and discard them from \mathcal{P}_s to get a point process \mathcal{P}^* . If $|\mathcal{P}_{s,x}| < j$ we set $\mathcal{P}^* := \mathcal{P}_s \cup \mathcal{Y}_x$, where we set

$$\mathcal{Y}_x := \{Y_1, \dots, Y_{j-|\mathcal{P}_{s,x}|}\}.$$

Let \mathcal{G}_s^* denote the subgraph of \mathcal{G}_s induced by $\mathcal{P}^* \cup \{x\}$. Then x has degree j in \mathcal{G}_s^* . Let V_x be the number of vertices in \mathcal{G}_s^* having degree j , other than x . This has the conditional distribution of $F_s(\{(x, \tau)\} \cup \eta_s) - 1$ given that $f_s(\{(x, \tau)\}, \eta_s) = 1$. This is because $\mathcal{P}_{s,x}$ and \mathcal{P}_s^x are independent Poisson processes, and conditioning on $f_s(\{(x, \tau)\}, \eta_s) = 1$ amounts to conditioning on the first of these two Poisson processes having j points.

If $|\mathcal{P}_{s,x}| > j$ then $|U_x - V_x| \leq U'_x + V'_x$, where we set U'_x to be the number of $y \in \mathcal{P}_{s,x}$ such that y has j neighbours in \mathcal{P}_s , and V'_x to be the number of pairs (y, z) with $y \in \mathcal{P}_{s,x}$, $z \in \mathcal{P}_s$, such that $z \neq y$, z is connected to y and z has at most j neighbours in \mathcal{P}_s^x . Then by the Mecke formula, and the assumption that $\phi_s \in \Phi_\varepsilon$ for all s ,

$$\begin{aligned} \mathbb{E} U'_x &= \int s\phi_s(x, y) \left(\frac{(s \int \phi_s(y, z)\mu(dz))^j}{j!} \right) \exp \left(- \int s\phi_s(y, z)\mu(dz) \right) \mu(dy) \\ &= O((sa_s)^{j+1}) \exp(-\Theta(sa_s)) = o(1), \end{aligned} \tag{5.7}$$

uniformly over $x \in \mathbb{X}$. Also, using that $\bar{\phi}_s \leq 1 - \varepsilon$ for all s , we have that

$$\begin{aligned} \mathbb{E} V'_x &= \sum_{i=0}^j \int s \phi_s(x, y) \int s \phi_s(y, z) \left(\frac{(s \int \phi_s(z, w)(1 - \phi_s(x, w))\mu(dw))^i}{i!} \right) \\ &\quad \times \exp \left(- \int s \phi_s(z, w)(1 - \phi_s(x, w))\mu(dw) \right) \mu(dz) \mu(dy) \\ &= O((sa_s)^{j+2}) \times \exp(-\Theta(sa_s)) \rightarrow 0, \end{aligned} \quad (5.8)$$

uniformly over $x \in \mathbb{X}$.

If $|\mathcal{P}_{s,x}| < j$ then $|U_x - V_x| \leq \tilde{U}_x + \tilde{V}_x$, where we set \tilde{U}_x to be the number of $y \in \mathcal{Y}_x$ having at most j neighbours in \mathcal{P}_s , and \tilde{V}_x is the number of pairs (y, z) with $y \in \mathcal{Y}_x$, $z \in \mathcal{P}_s$ connected to y , and z having at most j neighbours in \mathcal{P}_s . Then

$$\mathbb{E} \tilde{U}_x \leq j \sum_{i=0}^j \int (sa_s)^i \exp(-\Theta(sa_s)) \left(\frac{\phi_s(x, y)}{\int \phi_s(x, w)\mu(dw)} \right) \mu(dy) \rightarrow 0, \quad (5.9)$$

uniformly over $x \in \mathbb{X}$. Also

$$\begin{aligned} \mathbb{E} \tilde{V}_x &\leq \sum_{i=0}^j \int s \left(s \int \phi_s(z, w)\mu(dw) \right)^i \exp \left(- \int s \phi_s(z, w)\mu(dw) \right) \\ &\quad \times j \int \left(\frac{\phi_s(x, y)\phi_s(z, y)}{\int \phi_s(x, w)\mu(dw)} \right) \mu(dy) \mu(dz) \\ &= O((sa_s)^{j+1}) \times \exp(-\Theta(sa_s)) \rightarrow 0, \end{aligned} \quad (5.10)$$

uniformly over $x \in \mathbb{X}$.

Combining the estimates (5.7), (5.8), (5.9) and (5.10), and using Theorem 3.1, gives us the first part of (2.1), namely $D_j(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathcal{D}} Z_\alpha$. It remains to prove the second part of (2.1), and of (2.2).

For $n \in \mathbb{N}$, consider $G(\mathcal{X}_n, \phi_n)$. Let $s(n) = n - n^{3/4}$ and $t(n) := n + n^{3/4}$. By Chebyshev's inequality, with high probability (i.e. with probability tending to 1) we have $Z_{s(n)} \leq n \leq Z_{t(n)}$ so that $\mathcal{P}_{s(n)} \subset \mathcal{X}_n \subset \mathcal{P}_{t(n)}$. Moreover, when this happens, $G(\mathcal{X}_n, \phi)$ is the subgraph of $G(\mathcal{P}_{t(n)}, \phi)$ induced by \mathcal{X}_n , and $G(\mathcal{P}_{s(n)}, \phi)$ is the subgraph of $G(\mathcal{X}_n, \phi)$ induced by $\mathcal{P}_{s(n)}$.

By (5.1), for $i \in \mathbb{N}_0$,

$$\left(\frac{s(n)}{n} \right)^{i+1} \leq \frac{ED_i(G(\mathcal{P}_{s(n)}, \phi_n))}{ED_i(G(\mathcal{P}_n, \phi_n))} \leq \left(\frac{s(n)}{n} \right)^{i+1} \sup_{x \in \mathbb{X}} \left(\exp \left(n^{3/4} \int \phi_n(x, y)\mu(dy) \right) \right)$$

and by (5.6), both the upper and the lower bound tend to 1. Therefore by (5.4) and (5.5) we have as $n \rightarrow \infty$ that

$$\mathbb{E} D_i(G(\mathcal{P}_{s(n)}, \phi_n)) \rightarrow 0, \quad i = 0, 1, \dots, j-1; \quad (5.11)$$

$$\mathbb{E} D_j(G(\mathcal{P}_{s(n)}, \phi_n)) \rightarrow \alpha. \quad (5.12)$$

Let $k \in \mathbb{N}_0$ with $k \leq j$. If $\mathcal{P}_{s(n)} \subset \mathcal{X}_n$, then

$$D_{\leq k}(G(\mathcal{X}_n, \phi_n)) - D_{\leq k}(G(\mathcal{P}_{s(n)}, \phi_n)) = S_n - R_n,$$

where S_n denotes the number of points of $\mathcal{X}_n \setminus \mathcal{P}_{s(n)}$ with degree at most k in $G(\mathcal{X}_n, \phi_n)$, and R_n is the number of points of $\mathcal{P}_{s(n)}$ with degree at most k in $G(\mathcal{P}_{s(n)}, \phi_n)$ but with degree at least $k+1$ in $G(\mathcal{X}_n, \phi_n)$.

Let S'_n denote the number of points of $\mathcal{P}_{t(n)} \setminus \mathcal{P}_{s(n)}$ that are connected to at most k points of $\mathcal{P}_{s(n)}$, and let R'_n be the number of points of $\mathcal{P}_{s(n)}$ with degree at most k in $G(\mathcal{P}_{s(n)}, \phi_n)$ but with degree at least $k+1$ in $G(\mathcal{P}_{t(n)}, \phi_n)$. If $\mathcal{P}_{s(n)} \subset \mathcal{X}_n \subset \mathcal{P}_{t(n)}$, then $S_n \leq S'_n$ and $R_n \leq R'_n$. Hence by Markov's inequality,

$$\mathbb{P}[D_{\leq k}(G(\mathcal{X}_n, \phi_n)) \neq D_{\leq k}(G(\mathcal{P}_{s(n)}, \phi_n))] \leq \mathbb{P}[\{Z_{s(n)} \leq n \leq Z_{t(n)}\}^c] + \mathbb{E} S'_n + \mathbb{E} R'_n.$$

Now

$$\begin{aligned} \mathbb{E} S'_n &= 2n^{3/4} \sum_{i=0}^k \int_{\mathbb{X}} \frac{(s(n) \int \phi_n(x, y) \mu(dy))^i}{i!} \exp\left(-s(n) \int \phi_n(x, y) dy\right) \mu(dx) \\ &= \left(\frac{2n^{3/4}}{s(n)}\right) \mathbb{E} D_{\leq k}(G(\mathcal{P}_{s(n)}, \phi_n)), \end{aligned}$$

which tends to zero by (5.11) and (5.12). Also

$$\begin{aligned} \mathbb{E} R'_n &\leq s(n) \sum_{i=0}^k \int \frac{(s(n) \int \phi_n(x, y) \mu(dy))^i}{i!} \\ &\quad \times \exp\left(-s(n) \int \phi_n(x, y) \mu(dy)\right) \left\{2n^{3/4} \int \phi_n(x, y) \mu(dy)\right\} \mu(dx) \\ &\leq 2n^{3/4} a_n \mathbb{E} D_{\leq k}(\mathcal{P}_{s(n)}, \phi_n), \end{aligned}$$

which tends to zero by (5.6), (5.11) and (5.12).

Therefore with high probability we have $D_{\leq k}(G(\mathcal{X}_n, \phi_n)) = D_{\leq k}(G(\mathcal{P}_{s(n)}, \phi_n))$. This holds both for $k = j$, and for $k = j - 1$. Hence using the first part of (2.1) and the first part of (2.2) we obtain the second part of (2.1) and of (2.2). \square

6 Proof of Theorem 2.2

Given $\phi \in \Phi$, given $k, \ell \in \mathbb{N}$, and given $\mathbf{x} = (x_1, \dots, x_k) \in \mathbb{X}^k$ and $\mathbf{y} = (y_1, \dots, y_\ell) \in \mathbb{X}^\ell$, set

$$\phi(\mathbf{x}, \mathbf{y}) := 1 - \prod_{i=1}^k \prod_{j=1}^\ell (1 - \phi(x_i, y_j)). \quad (6.1)$$

We also write $\phi(\{x_1, \dots, x_k\}, \{y_1, \dots, y_\ell\})$ for $\phi(\mathbf{x}, \mathbf{y})$ (allowing multiplicities in the sets $\{x_1, \dots, x_k\}$ and $\{y_1, \dots, y_\ell\}$); it is the probability that there is at least one edge in the random graph $G(\{x_1, \dots, x_k, y_1, \dots, y_\ell\}, \phi)$ connecting one of the vertices x_i to one of the vertices y_j . If $k = 1$ we write $\phi(x_1, \{y_1, \dots, y_\ell\})$ for $\phi(\{x_1\}, \{y_1, \dots, y_\ell\})$. Also, let $h_\phi(\mathbf{x})$ or $h_\phi(x_1, \dots, x_k)$ denote the probability that $G(\{x_1, \dots, x_k\}; \phi)$ is connected, or more precisely

$$h_\phi(x_1, \dots, x_k) = \sum_{\Gamma} \prod_{\{\{i,j\}:\{i,j\} \in E(\Gamma)\}} \phi(x_i, x_j) \prod_{\{\{i,j\}:i,j \notin E(\Gamma)\}} (1 - \phi(x_i, x_j)), \quad (6.2)$$

where the sum is over all connected graphs Γ on vertex set $\{1, \dots, k\}$, and $E(G)$ denotes the set of edges of a graph G . By the Mecke formula (3.5), and the equality in distribution of $G(\mathcal{P}_s, \phi)$ and $G_\phi(\eta_s)$ as discussed in Section 4,

$$\begin{aligned} & \mathbb{E} N_k(G(\mathcal{P}_s, \phi)) = \mathbb{E} N_k(G_\phi(\eta_s)) \\ &= \frac{s^k}{k!} \int_{\mathbb{X}^k} h_\phi(x_1, \dots, x_k) \exp \left(-s \int \phi(z; \{x_1, \dots, x_k\}) \mu(dz) \right) \mu^k(d(x_1, \dots, x_k)). \end{aligned} \quad (6.3)$$

Now fix $k \in \mathbb{N}$ and $\alpha \in (0, \infty)$. Assume throughout this section that $\phi_s \in \Phi$ for $s > 0$, and (unless explicitly stated otherwise) that

$$\lim_{s \rightarrow \infty} \mathbb{E} N_k(G(\mathcal{P}_s, \phi_s)) = \alpha \in (0, \infty). \quad (6.4)$$

Proof of Theorem 2.2 (a) (i). Assume that $\bar{\phi}_s = o(1/\log s)$. We shall use the method of moments.

For $n, \ell \in \mathbb{N}$ we write $(n)_\ell$ for the descending factorial $n(n-1) \cdots (n-\ell+1)$. Then $(N_k(G(\mathcal{P}_s, \phi_s)))_\ell$ is the number of ordered ℓ -tuples of distinct k -components of $G(\mathcal{P}_s, \phi_s)$. This equals the sum over all $k\ell$ -tuples $x_{1,1}, \dots, x_{1,k}, \dots, x_{\ell,1}, \dots, x_{\ell,k}$ of distinct points of \mathcal{P}_s , of the indicator of the event that for each $i \leq \ell$ the subgraph of $G(\mathcal{P}_s, \phi_s)$ induced by $x_{i,1}, \dots, x_{i,k}$ is connected and these vertices are not connected to any other vertices of $G(\mathcal{P}_s, \phi_s)$, divided by $(k!)^\ell$. Hence by the Mecke formula,

$$\begin{aligned} \mathbb{E} [(N_k(G(\mathcal{P}_s, \phi_s)))_\ell] &= \frac{s^{k\ell}}{(k!)^\ell} \int_{\mathbb{X}^{k\ell}} \left(\prod_{i=1}^{\ell} h_{\phi_s}(x_{i,1}, \dots, x_{i,k}) \right) u_s(x_{1,1}, \dots, x_{\ell,k}) \\ &\quad \times \exp \left(-s \int \phi_s(z, \{x_{1,1}, \dots, x_{\ell,k}\}) \mu(dz) \right) \mu^{k\ell}(d(x_{1,1}, \dots, x_{\ell,k})), \end{aligned} \quad (6.5)$$

where we set $u_s(\{x_{1,1}, \dots, x_{\ell,k}\})$ to be the probability that the graph $G(\{x_{1,1}, \dots, x_{\ell,k}\}, \phi_s)$ has no edge between any $x_{i,j}$ and $x_{i',j'}$ such that $i \neq i'$, that is,

$$u_s(x_{1,1}, \dots, x_{\ell,k}) := \prod_{(i_1, j_1), (i_2, j_2) \in [\ell] \times [k]: i_1 < i_2} (1 - \phi_s(x_{i_1, j_1}, x_{i_2, j_2})). \quad (6.6)$$

By our condition on $\bar{\phi}_s$ the value of $u_s(x_{1,1}, \dots, x_{\ell,k})$ tends to 1, uniformly over $(x_{1,1}, \dots, x_{\ell,k})$. Also, by the union bound

$$\phi_s(z, \{x_{1,1}, \dots, x_{\ell,k}\}) \leq \sum_{i=1}^{\ell} \phi_s(z, \{x_{i,1}, \dots, x_{i,k}\}).$$

Therefore by (6.5), writing just N_k for $N_k(G(\mathcal{P}_s, \phi_s))$, we have

$$\begin{aligned} \mathbb{E}[(N_k)_\ell] &\geq (1 + o(1)) \left(\frac{s^k}{k!} \int_{\mathbb{X}^k} h_{\phi_s}(x_1, \dots, x_k) \right. \\ &\times \exp \left(-s \int \phi_s(z, \{x_1, \dots, x_k\}) \mu(dz) \right) \mu^k(d(x_1, \dots, x_k)) \Big)^{\ell}. \end{aligned}$$

By (6.4) and (6.3), this lower bound for $\mathbb{E}[(N_k)_\ell]$ tends to α^ℓ as $s \rightarrow \infty$.

By the Bonferroni bound and the union bound we have for $(z, x_{1,1}, \dots, x_{\ell,k}) \in \mathbb{X}^{1+\ell k}$ that

$$\begin{aligned} \phi_s(z, \{x_{1,1}, \dots, x_{\ell,k}\}) &\geq \sum_{i=1}^{\ell} \phi_s(z, \{x_{i,1}, \dots, x_{i,k}\}) \left(1 - \sum_{j=i+1}^{\ell} \phi_s(z, \{x_{j,1}, \dots, x_{j,k}\}) \right) \\ &\geq \sum_{i=1}^{\ell} \phi_s(z, \{x_{i,1}, \dots, x_{i,k}\}) (1 - k\ell\bar{\phi}_s). \end{aligned} \quad (6.7)$$

Therefore by (6.5),

$$\begin{aligned} \mathbb{E}[(N_k)_\ell] &\leq \frac{s^{k\ell}}{(k!)^\ell} \int_{\mathbb{X}^{k\ell}} \left(\prod_{i=1}^{\ell} h_{\phi_s}(x_{i,1}, \dots, x_{i,k}) \right) \\ &\times \exp \left(-s \sum_{i=1}^{\ell} \int \phi_s(z, \{x_{i,1}, \dots, x_{i,k}\}) (1 - k\ell\bar{\phi}_s) \mu(dz) \right) \mu^{k\ell}(d(x_{1,1}, \dots, x_{\ell,k})) \\ &= \left(\frac{s^k}{k!} \int_{\mathbb{X}^k} h_{\phi_s}(x_1, \dots, x_k) \right. \\ &\times \exp \left(-s \int \phi_s(z, \{x_1, \dots, x_k\}) (1 - k\ell\bar{\phi}_s) \mu(dz) \right) \mu^k(d(x_1, \dots, x_k)) \Big)^{\ell}. \end{aligned}$$

Since μ is a probability measure and $x^{1-k\ell\bar{\phi}_s}$ is a concave function on $x \geq 0$, we

obtain by Jensen's inequality that

$$\begin{aligned}
\mathbb{E}[(N_k)_\ell] &\leq \frac{s^{k^2\ell^2\bar{\phi}_s}}{k!\ell^2k\bar{\phi}_s} \left(\frac{s^k}{k!} \int (h_{\phi_s}(x_1, \dots, x_k))^{1/(1-k\ell\bar{\phi}_s)} \right. \\
&\times \exp \left(-s \int \phi_s(x, \{x_1, \dots, x_k\}) \mu(dx) \right) \mu^k(d(x_1, \dots, x_k)) \Big)^{\ell(1-k\ell\bar{\phi}_s)} \\
&\leq \frac{s^{k^2\ell^2\bar{\phi}_s}}{k!\ell^2k\bar{\phi}_s} \left(\frac{s^k}{k!} \int h_{\phi_s}(x_1, \dots, x_k) \right. \\
&\times \exp \left(-s \int \phi_s(x, \{x_1, \dots, x_k\}) \mu(dx) \right) \mu^k(d(x_1, \dots, x_k)) \Big)^{\ell(1-k\ell\bar{\phi}_s)}
\end{aligned}$$

which tends to α^ℓ by (6.3), (6.4) and our assumption on $\bar{\phi}_s$. Thus $\mathbb{E}[(N_k)_\ell] \rightarrow \alpha^\ell$, so by the method of moments (see e.g. Theorem 1.22 of [3]) the result (i) follows. \square

Proof of Theorem 2.2 (a) (ii). Assume now that $\bar{\phi}_s = o(s^{-1/2})$. Again we use the method of moments. Write n for $\lfloor s \rfloor$, and set $N'_k = N_k(G(\mathcal{X}_n, \phi_n))$. Then by (6.6) and the union bound,

$$\begin{aligned}
\mathbb{E}[(N'_k)_\ell] &= \frac{(n)_{k\ell}}{k!\ell} \int \cdots \int \left(\prod_{i=1}^{\ell} h_{\phi_n}(x_{i,1}, \dots, x_{i,k}) \right) u_n(\{x_{i,1}, \dots, x_{\ell,k}\}) \\
&\times \left(1 - \int \phi_n(z; \{x_{1,1}, \dots, x_{\ell,k}\}) \mu(dz) \right)^{n-k\ell} \mu(dx_{1,1}) \cdots \mu(dx_{\ell,k}) \\
&\geq (1 + o(1)) \left(\frac{n^k}{k!} \right)^\ell \int \cdots \int \left(\prod_{i=1}^{\ell} h_{\phi_n}(x_{i,1}, \dots, x_{i,k}) \right) \\
&\times \left(1 - \sum_{i=1}^{\ell} \int \phi_n(z, \{x_{i,1}, \dots, x_{i,k}\}) \mu(dz) \right)^n \mu(dx_{1,1}) \cdots \mu(dx_{\ell,k}).
\end{aligned}$$

Using the bound $1 - x \geq \exp(-x - x^2)$ for small positive x , we have for large n that

$$\begin{aligned}
\mathbb{E}[(N'_k)_\ell] &\geq (1 + o(1)) \left(\frac{n^k}{k!} \right)^\ell \int \cdots \int \left(\prod_{i=1}^{\ell} h_{\phi_n}(x_{i,1}, \dots, x_{i,k}) \right) \\
&\times \exp \left(-n \left(\sum_{i=1}^{\ell} \int \phi_n(z, \{x_{i,1}, \dots, x_{i,k}\}) \mu(dz) \right) \right. \\
&\left. - n \left(\sum_{i=1}^{\ell} \sum_{j=1}^k \int \phi_n(z, x_{i,j}) \mu(dz) \right)^2 \right) \mu(dx_{1,1}) \cdots \mu(dx_{\ell,k}),
\end{aligned}$$

so that

$$\begin{aligned} \mathbb{E}[(N'_k)_\ell] &\geq (1 + o(1)) \left(\frac{n^k}{k!} \int h_{\phi_n}(x_1, \dots, x_k) \right. \\ &\times \exp \left(-n \int \phi_n(z, \{x_1, \dots, x_k\}) \mu(dz) \right) \mu^k(d(x_1, \dots, x_k)) \Big)^\ell \times \exp(-nk^2 \ell^2 \bar{\phi}_n^2) \end{aligned}$$

which tends to α^ℓ by (6.3), (6.4), and the assumption that $\bar{\phi}_n = o(n^{-1/2})$.

Conversely, by the bound $e^x \geq 1 + x$ for all $x \in \mathbb{R}$, we also have

$$\begin{aligned} \mathbb{E}[(N'_k)_\ell] &\leq \left(\frac{n^k}{k!} \right)^\ell \int \cdots \int \left(\prod_{i=1}^\ell h_{\phi_n}(x_{i,1}, \dots, x_{i,k}) \right) \\ &\times \exp \left(-(n - \ell k) \int \phi_n(z; \{x_{1,1}, \dots, x_{\ell,k}\}) \mu(dz) \right) \mu(dx_{1,1}) \cdots \mu(dx_{\ell,k}) \end{aligned}$$

and using the Bonferroni bound as in (6.7), we obtain that

$$\begin{aligned} \mathbb{E}[(N'_k)_\ell] &\leq e^{k^2 \ell^2 \bar{\phi}_n} \left(\frac{n^k}{k!} \right)^\ell \int \cdots \int \left(\prod_{i=1}^\ell h_{\phi_n}(x_{i,1}, \dots, x_{i,k}) \right) \\ &\times \exp \left(-n \sum_{i=1}^\ell \int \phi_n(z; \{x_{i,1}, \dots, x_{i,k}\}) (1 - k\ell \bar{\phi}_n) \mu(dz) \right) \mu(dx_{1,1}) \cdots \mu(dx_{\ell,k}). \end{aligned}$$

By Jensen's inequality, since $x^{1-k\ell\bar{\phi}_n}$ is a concave function on $x \geq 0$, we have

$$\begin{aligned} \mathbb{E}[(N'_k)_\ell] &\leq e^{k^2 \ell^2 \bar{\phi}_n} \left(\frac{n^k}{k!} \right)^\ell \left(\int \cdots \int \left(\prod_{i=1}^\ell h_{\phi_n}(x_{i,1}, \dots, x_{i,k}) \right)^{1/(1-k\ell\bar{\phi}_n)} \right. \\ &\times \exp \left(-n \sum_{i=1}^\ell \int \phi_n(z; \{x_{i,1}, \dots, x_{i,k}\}) \mu(dz) \right) \mu(dx_{1,1}) \cdots \mu(dx_{\ell,k}) \Big)^{1-k\ell\bar{\phi}_n} \\ &\leq e^{k^2 \ell^2 \bar{\phi}_n} n^{k^2 \ell^2 \bar{\phi}_n} \left(\frac{1}{k!} \right)^\ell \left(n^k \int \cdots \int h_{\phi_n}(x_1, \dots, x_k) \right. \\ &\times \exp \left(-n \int \phi_n(z; \{x_1, \dots, x_k\}) \mu(dz) \right) \mu(dx_1) \cdots \mu(dx_k) \Big)^{\ell(1-k\ell\bar{\phi}_n)}. \end{aligned}$$

Hence, since our assumption $\bar{\phi}_n = o(n^{-1/2})$ implies that $\bar{\phi}_n = o(1/(\log n))$, we have $\limsup \mathbb{E}[(N'_k)_\ell] \leq \alpha^\ell$. Therefore $\mathbb{E}[(N'_k)_\ell] \rightarrow \alpha^\ell$, and the method of moments gives us part (ii). \square

Proof of Theorem 2.2 (a) (iii) and (b). Assume there exists $\varepsilon > 0$ such that $\phi_s \in \Phi_\varepsilon$ and $\bar{\phi}_s \leq 1 - \varepsilon$ for all s . Set $a_s = \sup_{x \in \mathbb{X}} \int \phi_s(x, y) \mu(dy)$. By (6.3),

$$\mathbb{E} N_k(G(\mathcal{P}_s, \phi_s)) = \Theta(s^k a_s^{k-1}) \exp(-\Theta(sa_s)). \quad (6.8)$$

If sa_s remains bounded away from zero and infinity, then by (6.8) we have that $\mathbb{E} N_k(G(\mathcal{P}_s, \phi_s)) \rightarrow \infty$, contradicting (6.4). Therefore for any sequence of values of s tending to infinity, there is a subsequence such that either $sa_s \rightarrow 0$ or $sa_s \rightarrow \infty$ as $s \rightarrow \infty$ along the subsequence.

Consider first the case with $sa_s \rightarrow 0$. In this case, by (6.8) and (6.4) we have $k \geq 2$ and $s^k a_s^{k-1} = \Theta(1)$. Recalling that $H_k(G)$ denotes the number of connected induced subgraphs of a graph G of order k , we have

$$\mathbb{E} H_{k+1}(G(\mathcal{P}_s, \phi_s)) = O(s^{k+1} a_s^k) = o(1).$$

Since $0 \leq H_k(G(\mathcal{P}_s, \phi_s)) - N_k(G(\mathcal{P}_s, \phi_s)) \leq (k+1)H_{k+1}(G(\mathcal{P}_s, \phi_s))$, we have

$$\mathbb{E} [H_k(G(\mathcal{P}_s, \phi_s)) - N_k(G(\mathcal{P}_s, \phi_s))] \rightarrow 0,$$

so by (6.4) we have that $\mathbb{E} H_k(G(\mathcal{P}_s, \phi_s)) \rightarrow \alpha$. Hence by Theorem 2.3 we have $H_k(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathcal{D}} Z_\alpha$. Also, by Markov's inequality, $\mathbb{P}[H_k(G(\mathcal{P}_s, \phi_s)) - N_k(G(\mathcal{P}_s, \phi_s)) \geq 1] \rightarrow 0$, so we also have $N_k(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathcal{D}} Z_\alpha$, which is the first part of (2.3).

Suppose now that $sa_s \rightarrow \infty$. There exists $K < \infty$ such that if $sa_s > K \log s$ then by (6.8),

$$\mathbb{E} N_k(G(\mathcal{P}_s, \phi_s)) = O(s(sa_s)^{k-1} \times e^{-\Theta(sa_s)} \times s^{-2})$$

which tends to zero, contradicting (6.4). But also if there exists $\delta > 0$ such that $sa_s < \delta \log s$, then there is a constant $\delta' > 0$ such that

$$\mathbb{E} N_k(G(\mathcal{P}_s, \phi_s)) \geq \delta' s(sa_s)^{k-1} s^{-1/2},$$

which tends to infinity, contradicting (6.4). Therefore we must have

$$sa_s = \Theta(\log s). \quad (6.9)$$

We seek to apply Theorem 3.1. For $s > 0$, let $\eta_s, Z_s, \tau, \tau_1, \tau_2, \dots$ be as in Section 4. Let \tilde{f}_s be the function \tilde{f} considered in Lemma 4.2, using the connection function $\phi \equiv \phi_s$. That is, let $\tilde{f}_s(x_1, \mathbf{t}_1, \dots, x_k, \mathbf{t}_k, \xi)$ be the indicator of the statement that $\{x_1, \dots, x_k\}$ induces a component of $G_{\phi_s}(\cup_{i=1}^k \{(x_i, \mathbf{t}_i)\} \cup \xi)$. Then with \tilde{F}_s denoting the function F obtained by taking $f \equiv \tilde{f}_s$ in the definition (3.2), we have that $\tilde{F}_s(\eta_s) = N_k(G_{\phi_s}(\eta_s))$ which has the same distribution as $N_k(G(\mathcal{P}_s, \phi_s))$.

For $\mathbf{x} = (x_1, \dots, x_k) \in \mathbb{X}^k$, set $\mathcal{X} = \{x_1, \dots, x_k\}$ (allowing multiplicities). Define the graph

$$\tilde{\mathcal{G}}_s := G_{\phi_s}(\eta_s \cup \{(x_1, \tau_1), \dots, (x_k, \tau_k)\}). \quad (6.10)$$

Let $\mathcal{P}_{s,\mathbf{x}}$ be the set of points of \mathcal{P}_s connected to at least one point of \mathcal{X} in $\tilde{\mathcal{G}}_s$ and let $\mathcal{P}_s^\mathbf{x} = \mathcal{P}_s \setminus \mathcal{P}_{s,\mathbf{x}}$.

The subgraph of $\tilde{\mathcal{G}}_s$ induced by vertex set \mathcal{P}_s has the same distribution as $G(\mathcal{P}_s, \phi_s)$, and we shall refer to this subgraph as $G'(\mathcal{P}_s, \phi_s)$. Likewise, we refer to the subgraph of $\tilde{\mathcal{G}}_s$ induced by vertex set $\mathcal{P}_s^\mathbf{x}$ as $G'(\mathcal{P}_s^\mathbf{x}, \phi_s)$, and we refer to the subgraph of $\tilde{\mathcal{G}}_s$ induced by vertex set \mathcal{X} as $G'(\mathcal{X}, \phi_s)$.

Let $U_\mathbf{x} = N_k(G'(\mathcal{P}_s, \phi_s))$. This has the same distribution as $\tilde{F}_s(\eta_s)$.

Let $V_\mathbf{x} = N_k(G'(\mathcal{P}_s^\mathbf{x}, \phi_s))$. We claim that this has the same distribution as conditional distribution of $\tilde{F}_s(\cup_{i=1}^k \{(x_i, \tau_i)\} \cup \eta_s) - 1$ given that $\tilde{f}_s(\cup_{i=1}^k \{(x_i, \tau_i)\}, \eta_s) = 1$. This is because by the Marking Theorem for Poisson processes (see e.g. [12]), the point processes $\mathcal{P}_{s,\mathbf{x}}$ and $\mathcal{P}_s^\mathbf{x}$ are independent and the statement that $\tilde{\mathcal{G}}_s$ has \mathcal{X} as the vertex set of a component is the equivalent to the statement that (i) $\mathcal{P}_{s,\mathbf{x}}$ has no points and (ii) the graph $G'(\mathcal{X}, \phi_s)$ is connected, which is independent of the outcome of $G'(\mathcal{P}_s^\mathbf{x}, \phi_s)$.

Also, $U_\mathbf{x} - V_\mathbf{x} = U'_\mathbf{x} - V'_\mathbf{x}$, where $U'_\mathbf{x}$ denotes the number of k -components of $G'(\mathcal{P}_s, \phi_s)$ with at least one vertex in $\mathcal{P}_{s,\mathbf{x}}$, and $V'_\mathbf{x}$ is the number of k -components of $G'(\mathcal{P}_s^\mathbf{x}, \phi_s)$ with at least one neighbour in $\mathcal{P}_{s,\mathbf{x}}$.

By the Mecke formula

$$\mathbb{E} U'_\mathbf{x} = \frac{s^k}{k!} \int_{\mathbb{X}^k} \phi_s(\mathbf{x}, \mathbf{y}) h_{\phi_s}(\mathbf{y}) \exp \left(-s \int \phi_s(z, \mathbf{y}) \mu(dz) \right) \mu^k(d\mathbf{y}),$$

where $\phi_s(\mathbf{x}, \mathbf{y})$ is given by (6.1). Then since we can choose the elements y_1, \dots, y_k of \mathbf{y} in an order such that y_1 is connected to \mathcal{X} and for each $j \geq 2$, y_j is connected to $\{y_1, \dots, y_{j-1}\}$, we have

$$\mathbb{E} U'_\mathbf{x} = O(s^k a_s^k) \times \exp(-\Theta(sa_s)), \quad (6.11)$$

which tends to zero (uniformly over \mathbf{x}).

Now $V'_\mathbf{x}$ is bounded by the number of pairs (y, \mathbf{z}) with $y \in \mathcal{P}_{s,\mathbf{x}}$ and $\mathbf{z} = (z_1, \dots, z_k)$ with $\{z_1, \dots, z_k\}$ inducing a k -component of $G'(\mathcal{P}_s^\mathbf{x}, \phi_n)$ and \mathbf{z} connected to y . Hence by the Mecke equation,

$$\begin{aligned} \mathbb{E} V'_\mathbf{x} &\leq s^{k+1} \int_{\mathbb{X}} \int_{\mathbb{X}^k} \phi_s(y, \mathbf{x}) \phi_s(y, \mathbf{z}) (1 - \phi_s(\mathbf{x}, \mathbf{z})) h_{\phi_s}(\mathbf{z}) \\ &\times \exp \left(-s \int \phi_s(w, \mathbf{z}) (1 - \phi_s(w, \mathbf{x})) \mu(dw) \right) \mu^k(d\mathbf{z}) \mu(dy) \\ &= O((sa_s)^{k+1}) \exp(-\Theta(sa_s)), \end{aligned} \quad (6.12)$$

which tends to zero, uniformly over \mathbf{x} ; here we have used the assumption that $\overline{\phi}_s \leq 1 - \varepsilon$, so that $\phi_s(w, \mathbf{z}) \leq 1 - \varepsilon^k$, for all s, w, \mathbf{z} . Therefore $\mathbb{E}[|U_{\mathbf{x}} - V_{\mathbf{x}}|] = \mathbb{E}[|U'_{\mathbf{x}} - V'_{\mathbf{x}}|] \rightarrow 0$, uniformly over $\mathbf{x} \in \mathbb{X}^k$. Then we can use Theorem 3.1 to get the first part of (2.3).

Before completing the proof of part (a) (iii) of Theorem 2.2, we prove part (b), so now instead of (6.4) we assume $\alpha_s := \mathbb{E} N_k(G(\mathcal{P}_s, \phi_s)) \rightarrow \infty$, but $\alpha_s = o(s)$. Then by (6.8), for every sequence of values of s tending to infinity, there is a subsequence such that either $sa_s \rightarrow \infty$ or $k \geq 2$ and $sa_s \rightarrow 0$ as $s \rightarrow \infty$ along the subsequence.

In both cases, the estimates (6.11) and (6.12) hold so by Theorem 3.1 we have $d_{TV}(\mathcal{L}(N_k(G(\mathcal{P}_s, \phi_s))), \mathcal{L}(Z_{\alpha_s})) \rightarrow 0$. Since also $(Z_{\alpha_s} - \alpha_s)/\sqrt{\alpha_s} \xrightarrow{\mathcal{D}} \mathcal{N}$, it follows that $(N_k(G(\mathcal{P}_s, \phi_s)) - \alpha_s)/\sqrt{\alpha_s} \xrightarrow{\mathcal{D}} \mathcal{N}$. This completes the proof of part (b) (ii).

Now we return to part (a), so we go back to assuming (6.4). As in the corresponding part of the proof of Theorem 2.1, for $n \in \mathbb{N}$ set $s(n) = n - n^{3/4}$ and $t(n) = n + n^{3/4}$. Then $\mathcal{P}_{s(n)} \subset \mathcal{X}_n \subset \mathcal{P}_{t(n)}$ with high probability, and also the point process $\mathcal{P}_{t(n)} \setminus \mathcal{P}_{s(n)}$ is a Poisson point process with mean measure $2n^{3/4}\mu(\cdot)$, independent of $\mathcal{P}_{s(n)}$. By (6.9),

$$\exp\left(n^{3/4} \int \phi_n(y, x) \mu(dy)\right) = \exp(n^{-1/4} \times \Theta(\log n)) = 1 + o(1),$$

uniformly over $x \in \mathbb{X}$, and therefore by (6.4) the sequence $(\phi_n)_{n \in \mathbb{N}}$ satisfies

$$\frac{s(n)^k}{k!} \int_{\mathbb{X}^k} \exp\left(-s(n) \int_{\mathbb{X}} \phi_n(z, \mathbf{x}) \mu(dz)\right) h_{\phi_n}(\mathbf{x}) \mu^k(d\mathbf{x}) \rightarrow \alpha. \quad (6.13)$$

For $n, \ell \in \mathbb{N}$ with $1 \leq \ell \leq k$, let $A_{n, \ell}$ be the event that at least one collection of ℓ of the added vertices of $\mathcal{X}_n \setminus \mathcal{P}_{s(n)}$ lies in a k -component of $G(\mathcal{X}_n, \phi_n)$. Let B_n be the event that at least one of the added vertices of $\mathcal{P}_{t(n)} \setminus \mathcal{P}_{s(n)}$ is connected to one of the k -components of $G(\mathcal{P}_{s(n)}, \phi_n)$.

If $A_{n, \ell}$ occurs and $\mathcal{P}_{s(n)} \subset \mathcal{X}_n \subset \mathcal{P}_{t(n)}$, then there is at least one pair $(\mathcal{X}, \mathcal{Y})$, such that $\mathcal{X} \subset \mathcal{P}_{s(n)}$ has $k - \ell$ elements, and $\mathcal{Y} \subset \mathcal{P}_{t(n)} \setminus \mathcal{P}_{s(n)}$ has ℓ elements, and $\mathcal{X} \cup \mathcal{Y}$ induces a connected subgraph of $G(\mathcal{P}_{t(n)}, \phi_n)$, and there is no connection between any vertex of $\mathcal{X} \cup \mathcal{Y}$ and any vertex of $\mathcal{P}_{s(n)} \setminus \mathcal{X}$ (however, we do allow other connections between vertices of $\mathcal{X} \cup \mathcal{Y}$ and other vertices of $\mathcal{P}_{t(n)} \setminus \mathcal{P}_{s(n)}$). By the Mecke equation, the expected number of such pairs equals

$$\begin{aligned} & \frac{(2n^{3/4})^\ell}{\ell!} \int \frac{(s(n))^{k-\ell}}{(k-\ell)!} \int h_{\phi_n}(x_1, \dots, x_{k-\ell}, y_1, \dots, y_\ell) \\ & \exp\left(-s(n) \int \phi_n(z, \{x_1, \dots, x_{k-\ell}, y_1, \dots, y_\ell\}) \mu(dz)\right) \\ & \mu^{k-\ell}(d(x_1, \dots, x_{k-\ell})) \mu^\ell(d(y_1, \dots, y_\ell)), \end{aligned}$$

which tends to zero by (6.13), so $\mathbb{P}[A_{n,\ell}] \rightarrow 0$. Also, the expected number of k -components in $G(\mathcal{P}_{s(n)}, \phi_n)$ which are connected to at least one vertex of $\mathcal{P}_{t(n)} \setminus \mathcal{P}_{s(n)}$ is at most

$$\frac{(s(n))^k}{k!} \int \exp \left(-s(n) \int \phi_n(z, \mathbf{x}) \mu(dz) \right) h_{\phi_n}(\mathbf{x}) 2n^{3/4} k a_n \mu^k(d\mathbf{x}),$$

and by (6.13) and (6.9) this tends to zero. Hence $\mathbb{P}[B_n] \rightarrow 0$. By the first part of (2.3) we have for $k \in \mathbb{N}$ that $N_k(G(\mathcal{P}_{s(n)}, \phi_n)) \xrightarrow{\mathcal{D}} Z_\alpha$. Also

$$\mathbb{P}[N_k(G(\mathcal{X}_n, \phi_n)) \neq N_k(G(\mathcal{P}_{s(n)}, \phi_n))] \leq \mathbb{P}[\cup_{\ell=1}^k A_{n,\ell}] + \mathbb{P}[B_n] + \mathbb{P}[\{Z_{s(n)} \leq n \leq Z_{t(n)}\}^c],$$

which tends to 0, and the second part of (2.3) follows. \square

7 Number of edges

For any graph G , according to our earlier notation $H_2(G)$ denotes the number of edges of G . Let $\phi \in \Phi$. Then

$$\mathbb{E} H_2(G(\mathcal{P}_s, \phi)) = \mathbb{E} H_2(G_{\phi_s}(\eta_s)) = \int_{\mathbb{X}} \int_{\mathbb{X}} \phi(x, y) s^2 \mu(dx) \mu(dy). \quad (7.1)$$

Theorem 7.1. *Suppose $\phi \in \Phi$. Set $\alpha := \mathbb{E} H_2(G(\mathcal{P}_s, \phi))$. Then*

$$d_{TV}(H_2(G(\mathcal{P}_s, \phi)), Z_\alpha) \leq (1 \wedge \alpha^{-1}) \int \left(\int \phi(x, y) s \mu(dy) \right)^2 s \mu(dx). \quad (7.2)$$

Proof. We shall use Theorem 3.1. Given $x, y \in \mathbb{X}$ with $\phi(x, y) > 0$, set $\tilde{\mathcal{G}}_s = G_{\phi_s}(\eta_s \cup \{(x, \tau_1), (y, \tau_2)\})$. Let $G'(\mathcal{P}_s, \phi_s)$ denote the subgraph of $\tilde{\mathcal{G}}_s$ induced by the vertex set \mathcal{P}_s , which has the same distribution as $G(\mathcal{P}_s, \phi_s)$.

Set $U_{x,y} = H_2(G'(\mathcal{P}_s, \phi_s))$ and $V_{x,y} = H_2(\tilde{\mathcal{G}}_s) - \mathbf{1}\{\{x, y\} \in E(\tilde{\mathcal{G}}_s)\}$, where $E(G)$ denotes the set of edges of a graph G . Then $V_{x,y}$ has the conditional distribution of $H_2(\tilde{\mathcal{G}}_s) - 1$ given that $\{x, y\} \in E(\tilde{\mathcal{G}}_s)$. Also $V_{x,y} \geq U_{x,y}$ and

$$\mathbb{E}[V_{x,y} - U_{x,y}] = \int_{\mathbb{X}} (\phi(x, z) + \phi(y, z)) s \mu(dz) =: w(x, y).$$

Hence by Theorem 3.1 the left hand side of (7.2) is bounded by the expression

$$\begin{aligned} & (1/2)(1 \wedge \alpha^{-1}) \int \int w(x, y) \phi(x, y) s^2 \mu(dx) \mu(dy) \\ &= (1 \wedge \alpha^{-1}) \int \left(\int \phi(x, y) s \mu(dy) \right)^2 s \mu(dx) \end{aligned}$$

as required. \square

For example, consider the geometric setting with μ having bounded, almost everywhere continuous density f with respect to Lebesgue measure on \mathbb{R}^d . Suppose $\phi_s(x, y) = \phi(r_s^{-1}(x - y))$ for some fixed integrable and almost everywhere continuous ϕ , and $r_s > 0$ satisfying $s^2 r_s^d \rightarrow \beta$ for some $\beta > 0$. Set $\alpha_s := \mathbb{E} H_2(G(\mathcal{P}_s, \phi_s))$. Then by (7.1),

$$\alpha_s \rightarrow \beta \left(\int \phi(z) dz \right) \int f(x)^2 dx =: \alpha.$$

By (7.2) we have $d_{TV}(\mathcal{L}(H_2(G(\mathcal{P}_s, \phi_s))), \mathcal{L}(Z_{\alpha_s})) = O(s^3 r_s^{2d}) = O(s^{-1})$.

In particular $H_2(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathcal{D}} Z_{\alpha}$. This could possibly also be proved by deriving a Poisson limit for $H_2(G(\mathcal{X}_n, \phi_n))$ by adapting the argument in [16, Theorem 3.4] to the RCM, and Poissonizing. However, the Poissonization would seem to introduce an error of at least $s^{-1/2}$ in the total variation distance, so the rate of convergence would not be as good.

Proof of Theorem 2.3. Let $k \in \mathbb{N}$ with $k \geq 2$. Let $\varepsilon > 0$ and assume $\phi_s \in \Phi_\varepsilon$ for all $s > 0$. Set $G_s := G(\mathcal{P}_s, \phi_s)$. Set $a_s = \sup_{x \in \mathbb{X}} \int \phi_s(x, y) \mu(dy)$. Then with $h_\phi(\cdot)$ defined at (6.2), we have

$$\mathbb{E} H_k(G_s) = \int s^k h_{\phi_s}(x_1, \dots, x_k) \mu^k(d(x_1, \dots, x_k)) = \Theta(s^k a_s^{k-1}), \quad (7.3)$$

because for each connected graph Γ on $\{1, \dots, k\}$ we can integrate the variables x_1, \dots, x_k in an order $(x_{\sigma(1)}, \dots, x_{\sigma(k)})$ such that for $2 \leq i \leq k$ each successive $\sigma(i)$ is connected in Γ to one or more of $\sigma(1), \dots, \sigma(i-1)$. Then each successive integral gives another factor of $\Theta(a_s)$.

Let $\mathbf{x} = (x_1, \dots, x_k) \in \mathbb{X}^k$, with x_1, \dots, x_k distinct and with $h_{\phi_s}(x_1, \dots, x_k) > 0$. Let $\tilde{\mathcal{G}}_s$ be the graph defined by (6.10), but now conditioned on the subgraph induced by $\{x_1, \dots, x_k\}$ being connected. Denote by $G'(\mathcal{P}_s, \phi_s)$ the subgraph of $\tilde{\mathcal{G}}_s$ induced by \mathcal{P}_s . Set $U_{\mathbf{x}} = H_k(G'(\mathcal{P}_s, \phi_s))$ and $V_{\mathbf{x}} = H_k(\tilde{\mathcal{G}}_s) - 1$. Then $U_{\mathbf{x}}$ has the distribution of $H_k(G_s)$ and $V_{\mathbf{x}}$ has the conditional distribution of $H_k(G(\mathcal{P}_s \cup \{x_1, \dots, x_k\}, \phi_s))$ given that $\{x_1, \dots, x_k\}$ induces a connected subgraph of this graph.

Now $V_{\mathbf{x}} \geq U_{\mathbf{x}}$ and we assert that

$$\mathbb{E} [V_{\mathbf{x}} - U_{\mathbf{x}}] = O \left(\sum_{j=1}^{k-1} (s^j a_s^j) \right), \quad (7.4)$$

uniformly over $\mathbf{x} \in \mathbb{X}^k$. To see this, observe that $V_{\mathbf{x}} - U_{\mathbf{x}}$ is the number of pairs $(\mathcal{X}, \mathcal{Y})$ with \mathcal{X} a non-empty subset of $\{x_1, \dots, x_k\}$ and \mathcal{Y} a non-empty subset of

\mathcal{P}_s , such that the subgraph of $\tilde{\mathcal{G}}_s$ induced by vertex set $\mathcal{X} \cup \mathcal{Y}$ is a connected graph of order k (to ease notation we ignore the issue of multiplicities in this notation). But then we can take the successive elements y_i of \mathcal{Y} in an order such that each of them is connected to at least one existing vertex from $\mathcal{X} \cup \{y_1, \dots, y_{i-1}\}$. Then each successive integral gives another factor of $O(sa_s)$.

For part (i), assume $\mathbb{E} H_k(G_s) \rightarrow \alpha \in (0, \infty)$. Then by (7.3) and (7.4) we have $\mathbb{E} |V_{\mathbf{x}} - U_{\mathbf{x}}| = o(1)$, uniformly over $\mathbf{x} \in \mathbb{X}^k$. Hence by Theorem 3.1, $H_k(G_s) \xrightarrow{\mathcal{D}} Z_\alpha$, which is part (i).

For part (ii), set $\alpha_s = \mathbb{E} H_k(G_s)$, and assume $\alpha_s \rightarrow \infty$ but $\alpha_s = o(s)$. Then by (7.3), the assumption $\alpha_s = o(s)$ implies that $sa_s \rightarrow 0$. Then by (7.4) and Theorem 3.1, we have that $d_{TV}(\mathcal{L}(H_k(G_s)), \mathcal{L}(Z_{\alpha_s})) \rightarrow 0$. Since also $(Z_{\alpha_s} - \alpha_s)/\sqrt{\alpha_s} \xrightarrow{\mathcal{D}} \mathcal{N}$, we therefore have $(H_k(G_s) - \alpha_s)/\sqrt{\alpha_s} \xrightarrow{\mathcal{D}} \mathcal{N}$. \square

8 Poisson U-statistics

Let $k \in \mathbb{N}$, and let $\mathbf{S}_k(\mathbb{X}) := \{\xi \in \mathbf{S}(\mathbb{X}) : |\xi| = k\}$. Let $h : \mathbf{S}_k(\mathbb{X}) \rightarrow \{0, 1\}$ be measurable. For $\xi \in \mathbf{S}(\mathbb{X})$ set

$$F(\xi) = \sum_{\psi \subset \xi : |\psi| = k} h(\psi).$$

Let $s > 0$, and let η be a Poisson process on \mathbb{X} with mean measure $\lambda := s\mu$. We seek to apply Theorem 3.1 to $W := F(\eta)$ for this choice of F , the so-called Poisson U -statistic. Assume μ is diffuse; as is the case elsewhere in this paper, this assumption is for notational convenience only.

For $\mathbf{x} = (x_1, \dots, x_k) \in \mathbb{X}^k$ with $h(\{x_1, \dots, x_k\}) = 1$ we set $U_{\mathbf{x}} = F(\eta)$, and

$$V_{\mathbf{x}} = F\left(\bigcup_{i=1}^k \{x_i\} \cup \eta\right) - 1.$$

Clearly $V_{\mathbf{x}}$ and $U_{\mathbf{x}}$ have the required distributional properties in the statement of Theorem 3.1. Also $V_{\mathbf{x}} \geq U_{\mathbf{x}}$, and

$$|V_{\mathbf{x}} - U_{\mathbf{x}}| = \sum_{J \subset [k], 1 \leq |J| < k} \sum_{\psi \subset \eta : |\psi| = k - |J|} h(\psi \cup \{x_i : i \in J\}),$$

so using the Mecke formula and integrating over \mathbf{x} we obtain that

$$\begin{aligned}
\int (\mathbb{E} |V_{\mathbf{x}} - U_{\mathbf{x}}|) h(\mathbf{x}) \lambda^k(d\mathbf{x}) &= \sum_{\ell=1}^{k-1} \frac{1}{(k-\ell)!} \binom{k}{\ell} \int \int h(\{x_1, \dots, x_\ell, y_1, \dots, y_{k-\ell}\}) \\
&\quad \times \lambda^{k-\ell}(d(y_1, \dots, y_{k-\ell})) h(x_1, \dots, x_\ell) \lambda^\ell(d\mathbf{x}) \\
&= \sum_{\ell=1}^{k-1} \frac{k!}{\ell!(k-\ell)!^2} \int \left(\int h(\{x_1, \dots, x_\ell, y_1, \dots, y_{k-\ell}\}) \lambda^{k-\ell}(d(y_1, \dots, y_{k-\ell})) \right)^2 \\
&\quad \times \lambda^\ell(d(x_1, \dots, x_\ell)) \\
&=: \gamma(h, \lambda).
\end{aligned}$$

By Theorem 3.1, if $\mathbb{E} F(\eta) = \alpha$, then $d_{TV}(\mathcal{L}(F(\eta)), \mathcal{L}(Z_\alpha)) \leq (1 \wedge \alpha^{-1}) \gamma(h, \lambda) / k!$. Also $d_W(\mathcal{L}(F(\eta)), \mathcal{L}(Z_\alpha)) \leq 3(1 \wedge \alpha^{-1/2}) \gamma(h, \lambda) / k!$. This bound is comparable to the one obtained in Theorem 7.1 of [5]. Our bound has an extra factor $1 \wedge \alpha^{-1/2}$ in front, which may make it better when α is large. Also, unlike [5] we do not make any topological assumptions on the measurable space \mathbb{X} . As remarked just after the statement of Theorem 3.1, it is possible to extend that result to the case where the measure λ is σ -finite, and hence to extend the above argument likewise, but we do not go into details here.

References

- [1] Barbour, A., Holst, L. and Janson, S. (1992). *Poisson Approximation*. Oxford Studies in Probability **2**, Oxford University Press, Oxford.
- [2] Bhamidi, S., van der Hofstad, R., and van Leeuwaarden, J. S. H. (2012). Novel scaling limits for critical inhomogeneous random graphs. *Ann. Probab.* **40**, 2299-2361.
- [3] Bollobás, B. (2001). *Random Graphs*. Second edition. *Cambridge Studies in Advanced Mathematics*, **73**. Cambridge University Press, Cambridge, 2001.
- [4] Bollobás, B., Janson, S. and Riordan, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31**, 3-122.
- [5] Decreusefond, L., Schulte, M. and Thäle C. (2015). Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry. *Ann. Probab.*, to appear.
- [6] Devroye, L. and Fraiman, N. (2014). Connectivity of inhomogeneous random graphs. *Random Structures Algorithms* **45**, 408-420.

- [7] Díaz, J., Petit, J., and Serna, M. (2000). Faulty random geometric networks. *Parallel Process. Lett.* **10**, 343-357.
- [8] Franceschetti, M. and Meester, R. (2007). *Random Networks for Communication*. Cambridge University Press, Cambridge.
- [9] Gupta, B. and Iyer, S. K. (2010). Criticality of the exponential rate of decay for the largest nearest-neighbor link in random geometric graphs. *Adv. in Appl. Probab.* **42**, 631-658.
- [10] Gupta, P. and Kumar, P. R. (1999). Critical power for asymptotic connectivity in wireless networks. *Stochastic Analysis, Control, Optimization and Applications*, 547-566, *Systems Control Found. Appl.*, Birkhäuser Boston, Boston.
- [11] Hsing, T. and Rootzén, H. (2005). Extremes on trees. *Ann. Probab.* **33**, 413-444.
- [12] Kingman, J. F. C. (1993). *Poisson Processes*. Oxford Studies in Probability **3**. Oxford University Press, Oxford.
- [13] Lindvall, T. (1992). *Lectures on the Coupling Method*. Wiley, New York.
- [14] Mao, G. and Anderson, B. D. O. (2012). Towards a better understanding of large-scale network models. *IEEE/ACM Transactions on Networking* **20**, 408-421.
- [15] Meester, R. and Roy, R. (1996). Continuum percolation. *Cambridge Tracts in Mathematics*, **119**. Cambridge University Press, Cambridge,
- [16] Penrose, M. (2003). *Random Geometric Graphs*. Oxford Studies in Probability **5**, Oxford University Press, Oxford.
- [17] Penrose, M. D. (2015). Connectivity of soft random geometric graphs. *Ann. Appl. Probab.*, to appear.
- [18] Penrose, M. D. (1998). Extremes for the minimal spanning tree on normally distributed points. *Adv. in Appl. Probab.* **30**, 628-639.
- [19] Penrose, M. D. (1999). On k -connectivity for a geometric random graph. *Random Structures Algorithms* **15**, 145-164.
- [20] Söderberg, B. (2002). General formalism for inhomogeneous random graphs. *Phys Rev. E* **66**, 06621
- [21] Yi, C.-W., Wan, P.-J., Lin, K.-W., and Huang, C.-H. (2006). Asymptotic distribution of the number of isolated nodes in wireless ad hoc networks with unreliable nodes and links. *Global Telecommunications Conference, 2006. GLOBECOM '06. IEEE*.